PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :		(11) International Publication Number:	**WO 98/13490**
C12N 15/12, C07K 14/705, C12Q 1/68, C12N 5/10, A01K 67/027, C12N 15/00, C12Q 1/02, G01N 33/68, C07K 16/18	A2	(43) International Publication Date:	2 April 1998 (02.04.98)

(21) International Application Number: PCT/NL97/00538

(22) International Filing Date: 26 September 1997 (26.09.97)

(30) Priority Data:
96202707.4   27 September 1996 (27.09.96)   EP
(34) Countries for which the regional or
    international application was filed: NL et al.

(71) Applicant (for all designated States except US): RIJKSUNI-VERSITEIT TE LEIDEN [NL/NL]; Stationsweg 46, NL–2312 AV Leiden (NL).

(72) Inventors; and
(75) Inventors/Applicants (for US only): FRANTS, Rune, Robert, Isak, Erik [NL/NL]; Wilsonstraat 85, NL–2131 PM Hoofddorp (NL). FERRARI, Michel, Dominique [NL/NL]; Pr. Bernardlaan 59, NL–2341 KJ Oegstgeest (NL). TERWINDT, Gisela, Marie [NL/NL]; Lanckhorstlaan 115, NL–2101 BC Heemstede (NL). OPHOFF, Roel, André [NL/NL]; Rijksuniversiteit te Leiden, Faculteit Geneeskunde, Wassenaarseweg 72, NL–2333 AL Leiden (NL).

(74) Agent: SMULDERS, Th., A., H., J.; Vereenigde Octrooibureaux, Nieuwe Parklaan 97, NL–2587 BN The Hague (NL).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: A GENE RELATED TO MIGRAINE IN MAN

(57) Abstract

    Genes for familial hemeplegic migraine (FHM), episodic ataxia type–2 (EA–2), common forms of migraine, and other episodic neurological disorders, such as epilepsy, have been mapped to chromosome 19p13 and chromosome 10p12. A brain–specific P/Q–type calcium channel subunit gene, covering 300 kb with 47 exons is provided. The exons and their surroundings reveal polymorphic variations and deleterious mutations that are linked to various types of calcium channel dysfunctions causing episodic neurological disorders in man or animals.

Title: A gene related to migraine in man.

Migraine is a frequent paroxysmal neuro-vascular
disorder, characterized by recurrent attacks of disabling
headache, vomiting, photo/phonophobia, malaise, and other
general symptoms (migraine without aura). Up to 20% of
5   patients may, in addition, experience transient neurological
(aura) symptoms during attacks (migraine with aura) (HCC,
1988). Up to 24% of females and 12% of males in the general
population are affected, however with variable attack
frequency, duration and severity (Russell et al., 1995).
10 . Knowledge about the mechanisms of the final common pathway of
migraine attacks has increased substantially the last five
years, resulting in improved, though still only symptomatic
(and sub-optimal) acute treatment for the attack. There is,
however, still very little knowledge about the etiology of
15   migraine attacks, i.e. why and how attacks begin and recur.
Accordingly, prophylactic treatment for migraine is non-
specific and has only limited efficacy.
Family, twin and population-based studies suggest that
genetic factors are involved in migraine, most likely as part
20   of a multifactorial mechanism (reviewed by Haan et al.,
1996). The complex genetics has hampered identification of
candidate genes for migraine. Familial Hemiplegic Migraine
(FHM) is a rare, autosomal dominant, subtype of migraine with
aura, associated with ictal hemiparesis and, in some families
25   cerebellar atrophy (HCC, 1988). Otherwise, the symptoms of
the headache and aura phase of FHM and "normal" migraine
attacks are very similar and both types of attacks may
alternate within subject and co-occur within families. FHM is
thus part of the migraine spectrum and can be used as a model
30   to study the complex genetics of the more common·forms of
migraine (Haan et al., 1996). A gene for FHM has been
assigned to chromosome 19p13 in about half of the families
tested (Joutel et al., 1993; Ophoff et al., 1994;

2

Joutel *et al.*, 1995). Remarkably, cerebellar atrophy was found only in families with FHM linked to chromosome 19p13, but not in unlinked families. Recently, we showed the 19p13 FHM locus to be also involved in "normal" migraine
5   (May *et al.*, 1995).

Episodic ataxia type 2 (EA-2) is another, autosomal dominant, paroxysmal neurological disorder, characterized by acetazolamide-responsive attacks of cerebellar ataxia and migraine-like symptoms, and interictal nystagmus and
10 cerebellar atrophy. Recently, a gene for EA-2 was assigned to chromosome 19p13, within the same interval as for FHM (Kramer *et al.*, 1995). This finding, as well as the clinical similarities, raise the possibility of EA-2 and FHM being allelic disorders.

15   Since other hereditary episodic neurological disorders responding to acetazolamide (such as hypokalaemic and hyperkalaemic periodic paralysis), as well as EA type-1 (which, in contrast to EA-2, is associated with continuous myokymia and non-responsive to acetazolamide) have all been
20 associated with mutations in genes encoding for ion channels (Ptacek *et al.*, 1991; Ptacek *et al.*, 1994; Brown *et al.*, 1994), we specifically looked for similar genes within the FHM and EA-2 candidate region.

In view of the above, the FHM/EA-2 locus can, since FHM
25 is part of the migraine spectrum, thus be used to study the genetic factors and biological mechanisms that are related to various episodic neurological disorders such as FHM, EA-2, common migraine and others such as epilepsy.

Calcium channels are multisubunit complexes composed of
30 at least an α1, an α2δ, and a β subunit. The central α1 subunit is functionally the most important component, acting as a voltage sensor and forming the ion-conducting pore. The other subunits have auxiliary regulatory roles. The membrane topology of the α1 subunit consist of four hydrophobic motifs
35 (I to IV), each containing six transmembrane α-helices (S1-

S6) and one hairpin (P) between S5-S6 that spans only the
outer part of the transmembrane region.

The present invention provides an isolated and/or
recombinant nucleic acid, or fragments thereof, encoding a
Ca^{2+}-channel α1 subunit related to familial hemiplegic
migraine and/or episodic ataxia type-2, derived from a gene
present on chromosome 19p13.1-19p13.2; a gene encoding the α1
(ion-conducting) subunit of a P/Q-type voltage gated calcium
channel. The present invention also provides access to and
methods to study the genetic background and identify other
subunits of the calcium channel subunit complexes and the
proteins related therewith that are associated with the
genetic factors and biological mechanisms that are related to
various episodic neurological disorders such as FHM, EA-2,
common migraine and others such as epilepsy which are related
to cation channel dysfunction.

The sequence of the cDNA of the gene is highly related
(≥90%) to a brain-specific rabbit and rat voltage gated
P/Q-type calcium channel α1 subunit (Mori *et al.*, 1991;
Starr *et al.*, 1991), and the open reading frame consists of
2261 amino acid residues. Northern blot analysis showed a
brain-specific expression, especially in the cerebellum.
Primary study of a cosmid contig harbouring the gene already
indicated an exon distribution over at least 300 kb of
genomic DNA. Recently, a neuronal Ca^{2+} α1A subunit gene was
localized to chromosome 19p13.1-p13.2 by FISH analysis
(Diriong *et al*, 1995). The gene symbol is CACNL1A4 and the
α1 subunit is classified as a P/Q-type. No sequence data for
the CACNL1A4 gene have been provided by Diriong or others,
but the same localization (chromosome 19p13.1) and the
identical classification (P/Q-type) suggests that the Ca^{2+}
channel α1 subunit we have identified is very similar to
CACNL1A4. No relation with migraine has been disclosed for
CACNL1A4. The genomic structures of three other human Ca^{2+}
channel α1 subunit genes (CACNL1A1, CACNL1A2 and CACNL1A3)
have been published to date (Hogan *et al*, 1994; Soldatov,
1994; Yamada *et al*, 1995). Both CACNL1A1 and CACNL1A2 span

about 150 kb and consist of 50 and 49 exons, respectively. The smaller CACNL1A3 gene is composed of 44 exons, distributed over 90 kb.

The present invention also provides an isolated and/or
5   recombinant nucleic acid comprising alleles of the invented gene which contain mutations relevant to the occurence of migraine and other neurological disorders which are related to cation channel dysfunction. Such mutations are for example a mutation at codon 192 resulting in the replacement of
10   arginine by glutamine (R192Q), and/or a mutation at codon 666 resulting in the replacement of threonine by methionine, and/or a mutation at codon 714 resulting in a replacement of valine by alanine and/or a mutation at codon 1811 resulting in a replacement of isoleucine by leucine, but also other
15   mutations of alleles of said gene which bear relationships with cation channnel dysfunction.

The present invention also provides isolated and/or recombinant nucleic acid comprising alleles of said gene which contain a polymorphic CA-repeat sequence specific for
20   various alleles of said gene. The present invention also provides isolated and/or recombinant nucleic acids comprising alleles of said gene which contain a CAG repeat.

The present invention also provides methods and tests (such as PCR, but also other tests to detect or amplify
25   nucleic acids are known in the art) to detect, identify and localize or distinguish genes and alleles of such genes, or fragments thereof, encoding for proteins or α, β or χ sub-units of specific *cerebral cation channels*, more specifically the invented gene and its various alleles encoding the α1
30   subunit of a P/Q-type voltage gated calcium channel and the gene encoding the β2 sub-unit, which are involved in the primary pathogenesis of neurological disorders such as FHM, migraine, EA-2 and SCA6. With such methods and tests one can study abnormalities of said gene.
35   The invention also provides recombinant expression vectors comprising isolated and/or recombinant nucleic acid comprising alleles of said genes or fragments therof, and

5

provides host cells or animals that comprise such vectors or
that are otherwise transformed with an isolated and/or
recombinant nucleic acid according to the invention.

5   The invention thus also provides a rationale and methods
for the testing and the development of specific prophylactic
medication for migraine and other episodic neurological, in
particular brain, disorders, such as epilepsy, associated
with cation channel dysfunction.

10   The invention for example provides cells or animals that
comprise recombinant vectors that comprise variants of said
genes or cells or animals that are transformed with said
variants. Also, the invention provides means to identify
naturally occuring variants of experimental animals with
changes in said gene related to FHM, EA-2, SCA7, migraine or

15   other neurological disorders associated with cation channel
dysfunction. An example of such an animal is the tottering
mouse, and its variants called leaner and rolling, described
in the experimental part of the invention. The invention also
provides cells or animals in which changes such as deletions

20   or mutations in said gene have been introduced by recombinant
nucleic acid techniques. All such cells or animals provided
by the invention can be used to study the pathophysiology of
FHM, EA-2, migraine or other neurological disorders
associated with cation channel dysfunction, for example to

25   test or develop specific medication for the treatment of said
disorders.

      The invention also provides proteins or peptides encoded
by said genes, or fragments thereof, related with cation
channel dysfunction, and detection of such proteins or

30   peptides by antibodies directed against said proteins or
peptides. Such antibodies can be of natural or synthetic
origin, and can be produced by methods known in the art. Such
proteins and antibodies and detection methods can be used to
further *in vitro* or *in vivo* studies towards the

35   pathophysiology of FHM, EA-2, migraine or other neurological
disorders associated with cation channel dysfunction, in
addition such proteins, antibodies and detection methods can

6

also be used to diagnose or identify such disorders in patients or in experimental animals.

Experimental Procedures

5

Subjects

Sixteen FHM patients were selected, including eight individuals from four unrelated chromosome 19-linked FHM families (*NL-A, UK-B, USA-C* (Ophoff *et al*, 1994), and *USA-P*

10 (Elliot *et al.*, 1995), two affected individuals from two small FHM families from Italy (*Italy I & II*) and six individuals with sporadic hemiplegic migraine (i.e. no other family member was shown to suffer from attacks of hemiplegic migraine). In families *NL-A* and *USA-P* cerebellar ataxia

15 and/or nystagmus is associated with FHM. An additional set of four subjects from four unrelated EA-2 families linked to chromosome 19, was also included (CAN-25, -45, -191, -197. Fifty randomly collected individuals from the Dutch population (Smith *et al.*, 1988) were used as a control to

20 determine the allele frequencies of polymorphic sites.

Patients with migraine with or without aura were diagnosed according to the international Headache Society (IHS) classification criteria. Patients attending the neurology outpatient clinic of Leiden University Medical

25 Center, The Netherlands, and patients responding to calls in local newspapers or in the periodical of the Dutch Migraine Patients Association, were screened for a positive family history of migraine. Only families with migraine in at least two generations were asked to participate. Probands (n=36)

30 and relatives (n=492) were personally examined and interviewed using semi-structured questionnaires. The questionnaire included questions about age at onset, frequency and duration of attacks, aura symtoms, premonitory signs and symptoms, triggers for attacks, medication, and

35 additional headaches. When family members were not available for a personal interview, information on their migraine was collected by interviewing their relatives. Because of the low

validity of diagnosing migraine auras through relatives, we
only assessed the presence or absence of migraine headaches.
Whenever possible, medical records were examined.

5   **Genomic structure**

Ten different cosmids from the contig extending the
invented gene, were subcloned separately in either M13 or
pBlueScript KS vector. From each cosmid library at least 3x96
random clones with an average insert size of about 2 kb, were
10   picked. Positive clones were identified by hybridization
techniques and subsequently sequenced with vector-specific
primers; intron-exon boundary sequences were completed using
cDNA-based primers.

15   **Mutation analysis, DHPLC and SSCP**

Genomic DNA was used as template to generate polymerase
chain reaction (PCR) products for single-strand
conformational polymorphism (SSCP) analysis and denaturing
high-performance liquid chromatography (DHPLC).
20   Amplifications were performed in standard conditions with
primer pairs as listed in Table 1 or listed below. Except for
the 5' side of exon 6, primers were chosen to produce
fragments that contained a single exon and at least 35
basepairs (including primer) of each flanking intron
25   sequence. Amplification of exons 1 and 20 was performed
producing two overlapping fragments and exon 19 was amplified
into three overlapping fragments. In addition, the following
markers;

		D10S191	Primer sequence 1	CTT TAA TTG CCC TGT CTT C
30			Primer sequence 2	TTA ATT CGA CCA CTT CCC
		D10S245	Primer sequence 1	AGT GAG ACT CGT CTC TAA TG
			Primer sequence 2	ACC TAC CTG AAT TCC TGA CC
35		D10S89	Primer sequence 1	AAC ACT AGT GAC ATT ATT TTC A
			Primer sequence 2	AGC TAG GCC TGA AGG CTT CT

8

DHPLC (Oefner et al., 1995; Hayward et al., 1996) was carried out on automated HPLC instrumentation. Crude PCR products, which had been subjected to an additional 3-minute 95°C denaturing step followed by gradual reannealing from

5   95-65°C over a period of 30 minutes prior to analysis, were eluted with a linear acetonitrile (9017-03, J.T. Baker, Phillipsburg, N.J., USA) gradient of 1.8% per minute at a flow-rate of 0.9 ml/min. The start- and end-points of the gradient were adjusted according to the size of the PCR

10   products (Huber et al., 1995). The temperature required for successful resolution of heteroduplex molecules was determined empirically by injecting one PCR product of each exon at increasing mobile phase temperatures until a significant decrease in retention was observed.

15   For SSCP analysis, primary PCR products were labeled by incorporation of [α-^{32}P]dCTP in a second round of PCR. Samples were diluted and denatured in formamide buffer before electrophoresis. SSCP was carried out according to published protocols (Orita et al., 1989; Glavac et al., 1994).

20   Digestion of several exons to yield products suitable for SSCP analysis.

         Sequencing of PCR products was performed with an ABI 377 automated sequencing apparatus with cycle sequencing according to the manufacturer. Furthermore, PCR products were

25   cloned in the TA vector (Invitrogen) and subjected to manual dideoxy sequence analysis (T7 Sequencing kit, Pharmacia Biotech.).

         A total of 481 blood samples were collected from patients with migraine. Genomic DNA was isolated as described

30   by Miller et al., 1988. The analyses of the highly informative microsatellite markers D19S391, D19S394, D19S221 and D19S226, D10S191, D10S245 and D10S89 were performed by PCR; primer sequences related to these markers are available through the human Genome Data Base (GDB).

35   The relative frequencies of marker alleles were estimated on the entire family material, with the relevant

correction for genetic relationships between individuals
(Boehnke, M, 1991) with the ILINK option of the I-INKAGE
package, version 5.03 (Lathrop *et al.*, 1985). The following
marker order and recombination frequencies were used in the
5   multipoint sib-pair analysis: D19S391-5%-D19S394-3%-D19S221-
5%-D19S226. Affected-sib-pair analysis was performed using
the MAPMAKER/SIBS software package, simultaneously including
marker information for all four DNA markers (Kruglyak, 1995).
Separate analyses were performed for migraine with aura,
10   migraine without aura, and a combination of both. Allowance
was made for dominance variance. When more than two affected
sibs occurred in a single sibship, weighted scores were
computed according to Suarez and Hodge (1979).
        In a sib-pair analysis, the occurrence of parental
15   marker alleles is compared among sibs. Normally, 25% of sib
pairs share their marker alleles from both parents, 50% share
one marker allele from one of their parents, while the
remaining 25% share no parental allele. Deviations from this
pattern towards increased sharing, and consistent with the
20   constraints of Holmans's (1993) possible triangle, are
explained as linkage (expressed as the maximum lod score
MLS). Increased sharing of marker alleles thus indicate that
the marker is located closely near a genetic risk factor. The
relative-risk ratio for a sib (λ_R), defined as the ratio of
25   the prevalence of a disease in sibs of affected individuals,
divided by the prevalence of a disease in the population, can
be calcutated (May *et al.*, 1995). In other words:

$$\lambda_R = \frac{\text{Affected risk for sib of a proband}}{\text{Affection risk for an individual in the general population}}$$

30

Results

35   **Genomic structure**
        The combination of hybridization and PCR strategies
resulted in a rapid assembly of the complete coding sequence

10

of the human cDNA, with an open reading frame of 6783
nucleotides encoding 2261 amino acid residues (figure 4). The
spatial distribution of the human Ca^{2+} channel expression was
assayed in rhesus monkey tissues. Total RNA was isolated from
5   several tissues, including various brain structures, and
probed with a human cDNA fragment. The probe detected a major
transcript of approximately 9.8 kb in cerebellum, cerebral
cortex, thalamus and hypothalamus, whereas no transcript was
detected in heart, kidney, liver or muscle. There was also no
10 hybridization signal found in RNA preparations from mouse
skin tissue or from human peripheral lymphocytes. In
addition, an attempt to amplify parts of the cDNA from human
peripheral lymphocytes failed.

    Complete alignment between the cDNA and individual exon
15 sequences was achieved, allowing the establishment of the
exon-intron structure (Table 1). The reconstruction of the
exon-intron structure of the CACNL1A4 gene revealed 47 exons
ranging in size from 36 bp (exon 44) to 810 bp (exon 19). The
exons are distributed over about 300 kb at genomic DNA level.
20 The result shows that the first 10 exons are located in a
region of about 150 kb covered by the first 5 cosmids of the
contig indicating relatively large introns at 5' side of the
gene. Sequences (Figure 1) were obtained of all exons
including approximately 100 bp of flanking introns, except
25 for intron 5 adjacent to exon 6. The forward primer of exon 6
harbours the splice junction and 3 bp of exon 6. Splice sites
around all exons are compatible with consensus sequence with
the exception of splice donor and acceptor of the first
intron.
30 The cosmid contig that yielded the initial Ca^{2+} channel
gene exons was extended to cover more than 300 kb.
Hybridization experiments showed that the first and last
cosmids of the contig were positive for 3'- and 5'-end cDNA
sequences, respectively, indicating a genomic distribution of
35 the gene over at least 300 kb (Figure 2). The cosmid contig
has been placed into the LLNL physical map of chromosome 19
at band p13.1, between the markers D19S221 and D19S226

11

(Figure 2). We identified a new polymorphic CA-repeat
sequence (D19S1150) on the cosmid contig. Oligonucleotide
primers (Table 1) flanking the repeat were synthesized and
amplification was performed by PCR as described. Analysis of
5   D19S1150 in 45 random individuals from the Dutch population
revealed nine alleles with an observed heterozygosity of
0.82. This highly polymorphic marker is located within the
gene and is therefore very useful in genetic analysis.

10   **Mutation analysis**

Exons and flanking intron sequences, containing the
complete coding region of CACNL1A4 and part of untranslated
sequences, were screened for the presence of mutations by
SSCP and DHPLC analysis in 20 individuals with either FHM or
15   EA-2. Several synonymous nucleotide substitutions and
polymorphisms were identified including a highly polymorphic
(CAG)n-repeat in the 3' untranslated region of exon 47
(Table 2). Of all polymorphisms only one was identified
predicting an amino acid change, an alanine to threonine
20   substitution at codon 454 (A454T).

Four different missense mutations were found in FHM
patients of which one mutation was observed in two unrelated
FHM affected individuals (Table 3). The mutations were shown
to segregate with the disease within the families; and were
25   not present in about 100 control chromosomes. A G-to-A
transition was identified in family *Italy-II* at codon 192,
resulting in a substitution of arginine to glutamine (R192Q)
within the first voltage sensor domain (IS4). A second
missense mutation occurs at codon 666, within the P-segment
30   of the second repeat replacing a threonine residue for
methione (T666M) in family *USA-P*. Two other mutations are
located in the 6th transmembrane spanning segment of
respectively repeat II and IV. The IIS6 mutation is a T-to-C
transition at codon 714, resulting in a substitution of
35   valine to alanine (V714A), identified in FHM family *UK-B*. The
mutation in domain IVS6 is an A-to-C transversion at codon
1811 resulting in a substitution of isoleucine to leucine

12

(I1811L). This I1811L mutation is found in family *NL-A* and family *USA-C*, two unrelated FHM families. Comparison of haplotypes in this region, including intragenic markers, did not reveal any evidence for a common founder of family *NL-A*

5   and *USA-C* (data not shown). No mutation was found in FHM family *Italy-I* nor in the six sporadic hemiplegic migraine patients. In addition to missense mutations in FHM families, we also identified mutations in two out of four EA-2 families (Table 3). In EA-2 family *CAN-191*, a basepair deletion occurs

10   in exxon 22 at nucleotide position 4073 causing a frameshift and a premature stop. The second EA-2 mutation is a transition of G-to-A of the first nucleotide of intron 24, predicted to leading to an aberrant splicing in family *CAN-26*. The invented gene also contains a CAG repeat,

15   of which expansions have been found in patients with autosomal dominant cerebellar ataxia (SCA6). Hence FHM, EA-2 and SCA6 are allelic ion channel disorders and different mutations are associated with different clinical symptomatologies.

20   Our study patients with common migraine (with or without aura) included 36 independent multigenerational Dutch families. At least some data were available on 937 family members and 212 persons who "married-in" (spouses). Of these, 442 family members (247 affected) and 86 spouses (7 affected)

25   were personally interviewed. The distribution of the different types of migraine among the 247 affected family members are as follows: 132 family members showed migraine without aura, 93 showed migraine with aura and 22 showed months-migraine, not fulfilling all critera by IHS. Among the

30   7 affected spouses these figures were 4, 1 and 2, respectively. We scored the parental transmission of migraine in the 36 families (Tabel 4) to investigate if an additional X-linked dominant or mitachondrial gene was involved. An approximately 2.5:1 preponderance of females among the

35   migraine sufferers was noted, which remained in the affected offspring. Affected fathers were found to transmit migraine

to their sons in 21 cases, making X-linked dominant or
mitochondrial inheritance highly unlikely.

The genetic analysis included 204 potentially affected
sib pairs; after correction for more than one sib pair in a
5   single sibship the total number of sib pairs was 108.
Affected-sib-pair analysis was performed for sib pairs who
were both affected with any form of migraine and, in separate
analyses, for sib pairs who where both suffering from either
migraine with aura or migraine without aura. The
10   informativeness of the region between the markers D19S391,
D19S394, D19S221 and D19S226 varied between 82% and 96%. The
combined analysis of migraine with and without aura resulted
in a maximum multipoint lod score of 1.69 (p≈0.005) with
marker D19S226. For migraine with aura the maximum multipoint
15   lod score was 1.29 corresponding with p≈0.013 with marker
D19S394. The maximum lod score for migraine without aura was
not significant (MLS <0.25) (data not shown). The relative
risk ratio for a sib to suffer from migraine with aura (λ_R),
defined as the increase in risk of the trait attributable to
20   the 19p13 locus, varied between λ_R=1.5 (for marker D19S394)
and λ_R=2.4 (for marker D19S226). When combining migraine with
and without aura, λ_R was 1.25. In a selected portion of 36
Dutch families with migraine with aura and without aura,
affected sib-pair analysis was performed for sib pairs who
25   were affected with any form of migraine. The following
markers, flanking the β2(CACNB2) calcium channel subunit gene
on chromosome 10p12, were tested: D108191, D10S246 and
D10S89. For the combined phenotype (migraine with and without
aura) a maximum pultipoint iod score of 0,95 (p<0,01) was
30   obtained with marker; D10S191. This result gives independent
evidence for a role of the P/Q type Ca^{2+} channel in migraine
and other neurological disorders.

Discussion

35   The genomic structure of the exemplified invented gene
revealed 47 exons distributed over about 300 kb (Table 1;
Figure 1). A comparison of the gene structure to already

14

known Ca^{2+} channel αl subunit genes (CACNL1A1, CACNL1A2, and CACNL1A3) (Soldatov, 1994; Yamada et al., 1995; Hogan et al., 1995), reveals a similar number of exons (50, 49, and 44 respectively) but a larger genomic span (300 kb vs

5   90-150 kb). Remarkebly, all splice sites are according to consensus sequence except for intron 1. Splice donor as well as splice acceptor of the first intron do not contain the expected gt...ag intron sequence. An incorrect cDNA sequence is unlikely because the cDNA sequence containing the junction

10   of the first two exons is identical to rabbit and rat sequence. Sequences corresponding to splice donor and acceptor are present in exon 1 and 2, suggesting an additional (yet unidentified) exon in the first intron encompassing part of sequences of exon 1 and exon 2.

15   To test the possible involvement of the invented gene relating to the CA^{2+}-channel sub-unit in migraine FHM, SCA6 and EA-2, we performed a mutation analysis by DHPLC and SSCP and found several alterations (For example Table 2 & 3). Only one missense variation was observed also present in 2% of the

20   normal controls (Table 2). This polymorphism is a alanine to threonine substitution at codon 454 (A454T), located in the intracellular loop between IS6 and IIS1 (Figure 2). This region contains a conserved alpha interaction domain (AID) that binds subunits (De Waard et al., 1996). However, A454T

25   is located outside the AID consensus sequence and is not likely to be involved.

          The identification of two mutations that disrupt the predicted translation product of the invented gene in two unrelated EA-2 patients and the segregation of these

30   mutations with the episodic ataxia phenotype in their families provide strong evidence that the invented gene is the EA-2 gene. A basepair deletion leads to a frame-shift in the putative translation product and encounters a stop codon in the next exon. The frame-shift in this EA-2 family is

35   predicted to yield a calcium channel αl subunit polypeptide consisting of repeat I and II, and a small portion of repeat III (IIIS1). The G-to-A transition of the first nucleotide of

15

intron 24 is affecting the nearly invariant GT dinucleotide of the intronic 5' splice junction. The brain-specific expression of the exemplified invented gene makes it extremely difficult to test the hypothesis that this mutation
5   produces aberrantly spliced RNAs by retaining the intron or utilizing other cryptic 5' splice sites.

The frameshift and splice site mutations in EA-2 may suggest a dominant negative effect of the truncated proteins by overruling the (corresponding) intact α1 subunits.

10   No mutations were found in the remaining EA-2 families (CAN-25 and -197). The use of two independent techniques for mutation screening (DHPLC and SSCP) makes it unlikely that we missed a heterozygote PCR product. Mutations in the promoter region or in intron sequences, resulting in aberrant

15   splicing, may have been the cause of EA-2 in these families. We could also have missed a mutation around the splice acceptor site of intron 5, covered by the forward primer of exon 6. However, larger deletions of e.g. complete exons with flanking intron sequence will disturb the predicted

20   translation product, like the ΔC_{4073} and splice site mutation do, but this is not detectable by a PCR-based screening method but can be seen Southern blot analysis instead.

Four different missense mutations were identified in five unrelated FHM families. These mutations all segregate

25   with FHM within a family and are not observed in over 100 normal chromosomes. The first missense mutation that we describe in the exemplified invented gene occurs in the IS4 domain of the α1 subunit (Table 3; Figure 3). The S4 domains are postulated to be voltage sensors because they have an

30   unusual pattern of positively charged residues at every third or fourth position separated by hydrophobic residues (Tanabe et al., 1987). In calcium channels the positively charged amino acid is an arginine residue (Stea et al., 1995). The mutation in FHM family *Italy-II* predicts a substitution of

35   the first arginine in the IS4 segment with a neutral, non-polar glutamine (R192Q). The change of the net positive

16

charge of this conserved region of the protein may influence
correct functioning of the voltage sensor.

The second missense mutation in FHM family *USA-P* occurs
in the P-segment of the second transmembrane repeat. A C-to-T
transition predicts substitution of a threonine residue with
methionine at codon 666 (T666M). Various observations have
shown that P-segments, the hairpin between S5 and S6 that
spans only the outer part of the transmembrane region, form
the ion-selectivity filter of the pore and binding sites for
toxins (Guy and Durell (1996). Alignment of protein sequence
of different P-segments indicating that some residues occur
in many different channel genes (Guy and Durell, 1996). The
T666M substitution alters one of the conserved residues in
the P-segment. It is conceivable that an alteration of a P-
segment affects the ion-selectivity or toxin binding of a
channel gene.

The remaining two missense mutations identified in FHM
families alter the S6 segment of the second and the fourth
repeat. A valine to alanine substitution in FHM family *UK-B*
is found in domain IIS6 at codon 714 (V714A). Domain IVS6 is
mutated in two unrelated FHM families (*NL-A* and *USA-C*),
predicting a isoleucine to leucine substitution at codon 1811
(I1811L). The V714A and I1811L missense mutations do not
really change the neutral-polar nature of the amino acid
residues. However, both S6 mutations are located nearly at
the same residue at the intracellular site of the segment and
are conserved in all calcium channel α1 subunit genes. In
addition, the A-to-C transversion leading the I1811L
substitution occurred in two unrelated FHM families on
different haplotypes indicating recurrent mutations rather
than a founder effect. Although the exact function of the S6
domains are not known, these data strongly suggest that
mutations in IIS6 and IVS6 result in FHM.

The I1811L mutation is present in two FHM families of
which one (*NL-A*) also displays a cerebellar atrophy in (some)
affected family members. The presence of cerebellar atrophy
in FHM families has been reported in about 40% of chromosome

17

19-linked FHM families, whereas none of the unlinked families was found to have cerebellar atrophy (Terwindt *et al.*, 1996).

The I1811L mutation excludes the possibility of allelic mutations in FHM and FHM with cerebellar atrophy. However, it is likely that FHM or FHM with cerebellar atrophy are the result of pleiotropic expression of a single defective gene.

No mutation was found in a small Italian FHM family (*Italy-I*). Apart from the possibilities discussed for EA-2, it should be noted that linkage to 19p13 was only suggested but never proved with significant lod scores (M. Ferrari, personal knowledge).

The four missense mutations identified indicate a mechanism for FHM in which both alleles of the $\alpha 1$ subunit are expressed, one harbouring an amino acid substitution which affects the function of this calcium channel $\alpha 1$ subunit by reducing or enhancing the electrical excitability. The relationship of FHM and other types of migraine makes it highly rewarding to investigate the involvement of the only missense variant observed (A454T) (Table 2), and to continue the search for other variants of the exemplified invented gene specific for common types of migraine.

The mutations in EA-2 and FHM demonstrate among others that the brain specific calcium channel gene CACNL1A4 is responsible for both EA-2 and FHM, and is also involved in the primary pathogenesis of the more common forms of migraine. We conducted the common migraine study in an independent sample of 36 extended Dutch families, with migraine with aura and migraine without aura. We found significant increased sharing of the marker alleles in sibs with migraine with aura (MLS=1.29 corresponding with $p \approx 0.013$). Although no such increased sharing was found for migraine without aura, a combined analysis for both migraine types resulted in an even more significant increased sharing (MLS=1.69 corresponding with $p \approx 0.005$). These results clearly indicate the involvement of the calcium α_{1A}-subunit gene region on 19p13 in both migraine with and without aura; the contribution to migraine with aura, however, seems strongest.

18

The positive findings in our study clearly demonstrate
an involvement of the FHM locus region in non-hemiplegic
familial migraine, notably in migraine with aura. The
P/Q-type calcium channel α_{1A}-subunit gene on chromosome
5   19p13 may be an "aura-gene" and is involved in both FHM and
migraine with aura, but not in migraine without aura. This
however, seems unlikely since an increased sharing of marker
alleles was also found when we combined the results for
migraine with and without aura. Furthermore, the increase in
10 sharing was stronger than expected to be only due to the
contribution of migraine with aura. An alternative
explanation is that the gene is involved in both types of
migraine, but in migraine without aura there is an additional
strong effect of other, possibly environmental factors,
15 thereby reducing the penetrance.

The latter view is also supported by the results
obtained from calculating the relative risk ratios (λ_R) for
sibs from affected individuals to also have migraine. The
relative risk ratio for a sib to suffer from migraine with
20 aura was $\lambda_R=2.4$. When combining migraine with and without
aura, λ_R was 1.25. In a population-based study the relative
risk for first degree relatives of probands with migraine
with aura to also have migraine with aura was $\lambda_R=3.8$. Because
of the female preponderance among migraine patients, X-linked
25 dominant or mitochondrial inheritance has been suggested to
be involved in familial migraine. Although a predominant
maternal inheritance pattern was noted in our families, X-
linked dominant or mitochondrial inheritance were found to be
highly unlikely because affected fathers transmit migraine to
30 their sons. Furthermore, the predominant maternal inheritance
can be explained by the female preponderance among the
migraine patients.

We conclude that the well-established genetic
contribution to the etiology of migraine is partly, but not
35 entirely, due to genetic factors located in the chromosomal
region of the P/Q-type calcium channel α_{1A}-subunit gene.
Further analysis of the cerebral distribution and function of

19

this calcium channel, as well as of the "mutated channels",
will help to unravel the pathogenetic pathway of migraine. It
may also contribute to a better understanding of the
mechanisms involved in related disorders such as episodic
5   ataxia type-2, autosomal dominant cerebellar ataxia (SCA6),
cerebral atrophy, and epilepsy, which all have been found to
be associated with mutations in this gene. Study of FHM, EA-2
mutants and variants such as the A454T variant expressed in
vitro or in mouse or other experimental animal models will
10   ultimately lead to better understanding of the diseases,
their cellular mechanisms, and the clinical relationship
between FHM, EA-2, migraine, and other episodic neurological
disorders such as epilepsy, and will provide rationales for
the development of prophylactic therapy.
15

**Localization and identification of the mouse gene related to
the neurological mouse mutations tottering, leaning and
rolling.**

The tottering (tg) mutation arose spontaneously in the
20   DBA inbred strain, and has been back-crossed into a C57BL/6J
(B6) inbred strain for at least 30 generations. The genome of
the tg mouse therefore is of B6 origin except for a small
region around the tg gene on chromosome 8. Interestingly, the
chromosome 8 region in mouse has synteny with the human
25   chromosome 19p13.1, in which the human calcium channel alpha1
subunit has been identified. We therefore consider the tg
locus as a possible site of the mouse homologue of the human
calcium channel gene.

To determine the exact localization of the mouse
30   homologue, PCR was carried out with primers based on human
cDNA sequence selected from Figure 1 and mouse genomic DNA as
template. In human, primers were known to be located in
different flanking exons. PCR amplification on human DNA
yielded a 1.5kb fragment.
35   Forward primer : 5'- caa cat cat gct ttc ctg cc-3'
Reversed primer: 5'- atg atg acg gcg aca aag ag-3'

20

Amplification on mouse DNA yielded a 750-bp fragment. The fragment mainly consists of intronic sequences. SSCP analysis revealed several polymorphisms in the different inbred strains (each strain a specific pattern). Analysis of

5 amplified product of the tg/tg (homozygote) and tg/+ (heterozygote) mice demonstrated a DBA specific signal in the tg/tg mouse, and a heterozygous pattern of DBA and B6 inbred strains in the heterozygous tg/+ mouse.

These results show that the mouse homologue of the human

10 calcium channel alpha1 subunit is located within the mouse tottering interval on chromosome 8.

In conclusion: the phenotypic characteristics of the tg mouse (tg/tg and tg/+) suggest involvement of ion-channels in the tg-etiology. The localization of the mouse homologue of

15 the human calcium gene within the tottering interval show that a tottering phenotype in mouse is caused by a mutation in the mouse homologue of the CACNL1A4 gene. With various variants of the tottering mouse (the Jackson Laboratory, Bar Habor, ME, USA), such as the leaner and rolling varieties,

20 such mutations in the mouse homologue of the CACNL1A4 gene can be found, clearly demonstrating that the gene is related to a variety of episodic neurologic disorders and using this genetic information one can engage in a variety of pathofysiological studies, as for example indicated below.

25 The tg mutation arose spontaneously in the DBA/2 inbred strain. tg/tg homozygotes are characterized by a wobbly gait affecting the hindquarters in particular, which begins at about 3 to 4 weeks of age, and by intermittent spontaneous seizures which resemble human epileptic absence seizures. The

30 central nervous system of the mice appears normal by light microscopy. There is no discernible cerebellar hypoplasia. In fluorescent histochemistry studies tg/tg mice show a marked increase in number of noradrenergic fibers in the terminal fields innervated by locus ceruleus axons, the hippocampus,

35 cerebellum, and dorsal lateral geniculate. Treatment of neonatal tg/tg mice with 6-hydroxydopamine, which selectively causes degeneration of distal noradrenergic axons from the

21

locus ceruleus, almost completely abolishes the ataxic and
seizure symptoms.

The leaner mutation of the tottering mouse arose
spontaneously in the AKR/J strain. Homozygotes are recognized
5   at 8 to 10 days of age by ataxia, stiffness, and retarded
motor activity. Adults are characterized by instability of
the trunk, and hypertonia of trunk and limb muscles. The
cerebellum is reduced in size, particularly in the anterior
region, in tg<la>/tg<la> mice, as is the case with a certaim
10   number of FHM patients. There is loss of granule cells
beginning at 10 days of age and loss of Purkinje and Golgi
cells beginning after 1 month. Cell loss later slows but
continues throughout life. Granule and Purkinje cells are
more severely affected than Golgi cells and the anterior
15   folia more severely affected than other parts of the
cerebellum. The cerebellum of tg<la>/tg mice shows shrinkage
and degenerative changes of the Purkinje cells. The loss in
cerebellar volume in tg<la>/tg and in tg/tg mice is specific
to the molecular layer, with no change in volume of the
20   granule cell layer or the white matter layer. Allelism of
leaner with tottering was shown in complementation and
linkage tests.
A third variety of the tottering mouse is (tg<rol>) called
the rolling Nagoya. Found among descendants of a cross
25   between the SIII and C57BL/6 strains, the tg<rol> mutation
apparently occurred in the SIII strain. Homozygotes show poor
motor coordination of hindlimbs that may lead to falling and
rolling, and sometimes show stiffness of the hindlimbs and
tail. No seizures have been observed. Symptoms are
30   recognizable at 10 to 14 days old. They appear a little
earlier than those of tg/tg mice and are somewhat more
severe. The cerebellum is grossly normal until 10 days of
age, but after that grows more slowly than normal. The size
of the anterior part of the central lobe of the cerebellum is
35   reduced with reduction in the numbers of granule, basket, and
stellate cells but normal numbers of Purkinje cells. There is
a reduced concentration of glutamate and an increasd

concentration of glycine and taurine in the cerebellum and decreased activity of tyrosine hydroxylase in the cerebellum and other areas.

23

Legends to figures

Figure 1
Nucleic acid sequences of 47 exons and flanking intron
5   sequences containing the complete coding region of the
invented gene and part of untranslated sequences.

Figure 2
Genetic map, cosmid contig and global exon distribution of
10   the invented gene om chromosome 19p13.1. The cosmid contog is
shown with EcoRI restriction sites, available via Lawrence
Livermore National Laboratory; exon positions are indicated
schematically, regardless of exon or intron sizes (Table 1).
D19S1150 is a highly polymorpmic intragenic (Ca)n-repeat.

15

Figure 3
Membrane topology of $\alpha 1$ subunit of the P/Q-type Ca^{2+}-channel.
The location and amino acid substitutions are indicated for
mutations that cause FHM or EA-2

20

Figure 4
The coding sequence of human cDNA of the invented gene with
an open reading frame encoding 2261 amino acid residues.

24

References

1.   Browne, D.L., Gancher S.T, Nutt, J.G., Brunt, E.R., Smith E.A., Kramer P., and Litt M. (1994). Episodic ataxia/myokymia syndrome is associated with point mutations in the human potassium channel gene, KCNA1. *Nat. Genet.* 8: 136-140.

2.   Diriong S., Lory P., Williams M.E., Ellis S.B., Harpold M.M., and Taviaux S. (1995). Chromosomal localization of the humangenesfor α1A, α1B, and α1E voltage-dependent Ca^{2+} channel subunits. *Genomics* 30: 605-609.

.   Headache Classification Committee (HCC) of the International Headache Society (1988). Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. *Cephalalgia* 8: 19-28.

4.   Hogan, K., Powers, P.A., and Gregg, R.G. (1994). Cloning of the human skeletal muscle alpha 1 subunit of the dihydropyridine-sensitive L-type calcium channel (CACNL1A3). *Genomics* 24: 608-609.

5.   Joutel A., Bousser M-G., Biousse V., Labauge P., Chabriat H., Nibbio A., Maciazek J., Meyer B., Bach M-A., Weissenbach J., Lathrop G.M., and Tournier-Lasserve E. (1993). A gene for familial hemiplegic migraine maps to chromosome 19. *Nature Genet.* 5: 40-45.

6.   Joutel A., Ducros A., Vahedi K., Labauge P., Delrieu O., Pinsard N., Mancini J., Ponsat G., Gaoftiere F., Gasant J.L., Maziaceck J, Weissenbach J., Bousser M.G., and Tournier-Lasserve E. (1994). Genetic heterogeneity of familial hemiplegic migraine. *Am. J. Hum. Genet.* 55: 1166-1172.

7.   Hayward-Lester, A., Chilton, B.S., Underhill, P.A., Oefner, P.J., Doris, P.A. (1996). Quantification of specific nucleic acids, regulated RNA processing and genomic polymorphisms using reversed-phase HPLC.
        In: F. Ferr (Ed.), Gene Quantification, Birkhuser Verlag, Basel, Switzerland.

25

8.   Huber, C.G., Oefner, P.J., Bonn, G.K. (1995) Rapid and accurate sizing of DNA fragments by ion-pair chromatography on alkylated nonporous poly(styrene-divinylbenzene) particles. *Anal. Chem.*, **67**, 578-585.

9.   Kramer P.L., Yue Q., Gancher S.T., Nutt J.G., Baloh R., Smith E., Browne D., Bussey K., Lovrien E., Nelson S, and Litt M. (1995). A locus for the nystagmus-associated form of episodic ataxia maps to an 11-cM region on chromosome 19p. *Am. J. Hum. Genet.* **57**: 182-185.

10.   May A, Ophoff R.A., Terwindt G.M., Urban C., Van Eijk R., Haan J., Diener H.C., Lindhout D., Frants R.R., Sandkuiji L.A., and Ferrari M.D. (1995). Familial hemiplegic migraine locus on 19p13 is involved in the common forms of migraine with and without aura. *Hum. Genet.* **96**: 604-608.

11.   Mori Y, Friedrich T., Kim M.S., Mikami A., Nakai J., Ruth P., Bosse E., Hofmann F., Flockerzi V., Furuichi T., Mikoshiba K. Imoto K., Tanabe T., and Numa S. (1991). Primary structure and functional expression from complementary DNA of a brain calcium channel. *Nature* **350**: 398-402.

12.   Oefner, P.J., Underhill, P.A. (1995) Comparative DNA sequencing by denaturing high-performance liquid chromatography (DHPLC). *Am. J. Hum. Genet.* 57 [Suppl.], A266.

13.   Ophoff R.A., Van Eijk R., Sandkuijl L.A., Terwindt G.M., Grubben C.P.M., Haan J., Lindhout D., Ferrari M.D., and Frants R.R. (1994). Genetic heterogeneity of familial hemiplegic migraine. *Genomics* **22**: 21-26.

14.   Orita, M., Suzuki, Y., Sekiya, T., and Hayashi, K. (1989). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5**: 874-879.

15.   Ptacek L.J., George A.L., Griggs R.C., Tawil R., Kallen R.G., Barchi R.L., Robertson M., and Leppert M.F. (1991). Identification of a mutation in the gene causing hyperkalemic periodic paralysis. *Cell* **67**: 1021-1027.

16.   Ptacek L.J., Tawil R., Griggs R.C., Engel A.G., Layzer R.B., Kwiecinski H., McManis P.G., Santiago L., Moore M., Fouad G., Bradley P., and Leppert M.F. (1994).

26

Dihydropyridine receptor mutations cause hypokalemic periodic paralysis. *Cell* **77**: 863-868.

17. Ravnik-Glavac, M., Glavac D., and Dean, M. (1994). Sensitivity of single-strand conformation polymorphism and heteroduplex method for mutation detection in the cystic fibrosis gene. *Hum. Mol. Genet.* **3**: 801-807.

18. Russell, M.B., Rasmussen B K., Thorvaldsen P., and Olesen J.. (1995). Prevalence and sex-ratio of the subtypes of migraine. *Int. J Epidemiol.* **24**: 612-618.

19. Starr T.V.B., Prystay W., and Snutch T. (1991). Primary structure of a calcium channel that is highly expressed in the rat cerebellum. *Proc. Natl. Acad. Sci.* **88**: 5621-5625.

20. Soldatov N.M. (1994) Genomic structure of Human L-type Ca^{2+} channel. *Genomics* **22**: 77-87.

21. Teh B.T., Silburn P., Lindblad K., Betz R., Boyle R., Schalling M., and Larsson C. (1995). Familial periodic cerebellar ataxia without myokymia maps to a 19-cM region on 19p13. *Am. J. Hum. Genet.* **56**: 1443-1449.

22. Terwindt G.M., Ophoff R.A., Haan J., Frants R.R., and Ferrari M.D. (1996). Familial hemiplegic migraine: a clinical comparison of families linked and unlinked to chromosome 19. *Cephalagia* **16**: 153-155.

23. Von Brederlow, B., Hahn, A.F., Koopman, W.J., Ebers, G.C., and Bulman, D. (1995). Mapping the gene for acetozolamide responsive hereditary paroxysmal cerebellar ataxia to chromosome 19p. *Hum. Mol. Genet.* **2**: 279-284.

24. Yamada Y., Masuda K., Li Q., Ihara Y, Kubota A., Miura T., Nakamura K., Fujii Y., Seino S., and Seino Y. (1995) The structures of the Human Calcium channel α1 subunit (CACNL1A2) and β-subunit (CACNLB3) genes. *Genomics* **27**: 312-319.

Exon/intron organization of the human invented gene
and exon-specific primer pairs

Table 1

Exon	cDNA	Size	Domain	Cosmid(s)	Primer Forward	Primer Reversed	Size
1	UTR - 568	:500		25960 / 30151	tct ccg cag tcg tag ctc ca cgc aaa gga tgt aca agc ag	ggt tgt aga gtg cca tgg tc att ccc aag cct cca ggg tag	320 370
2	569 - 674	106	I S1	30151	cac ctc caa cac cct tct tt	tct gtg ccc tgc tcc act c	240
3	675 - 814	140	I S2. I S3	30151	acg ctg acc ttg cct tct ct	caa cca aaa gcc tcg taa tc	230
4	815 - 906	92	I S3. I S4	28913	aaa acc cac cct ctg ttc tc	ttg tca ggg tcg gaa act ca	160
5	907 - 1059	153	I S5	28913 / 27415	ctt ggt ggc ggg gtt t	ctg cct aat cct ccc aag ag	290
6	1060 - 1253	194		27415	tcc ctt ccc ttt tgt aga tg	gtg ggg ctg tgt tgt cct t	350
7	1254 - 1357	104	I S6	27415	gac aga gcc aca aga gaa cc	agc aaa gag gag tga gtg gg	250
8	1358 - 1473	116		34077 / 27415	ata ctc tgg ctt ttc tat gc	gca tga ctc tct ttg tac tc	230
9	1474 - 1530	57		34077	gca gag aat ggg ggt gg	ctg agg tgg gtt tag agc ag	180
10	1531 - 1623	93		34077	ggg taa cgt ctt ttt ctc ttg c	atg tct ctt ggg cga tag gt	200
11	1624 - 1833	210	II S1	16894 / 32236	att tct tct gaa gga aca gc	gga ggg atc agg gag ttg gc	310
12	1834 - 1946	113	II S2. II S3	16894	caa gcc taa cct cct ctc tg	tca ttc cag gca aga gct g	200
13	1947 - 2051	105	II S3. II S4	16894	att tgg agg gag gag ttt gg	tca ctt tcc caa ctt tct gg	310
14	2052 - 2191	140	II S4. II S5	16894	cag aaa gtt ggg aaa gta gc	ttg aat tcc tgt gaa gga c	250
15	2192 - 2264	73		16894	ctt gga gat gag ata ctg agc	cag gca ctt tca tct gtg ac	200
16	2265 - 2382	118	II S6	16894	tcc aca gct gca tct cca ag	acc ctc cct tga gcc cct	270
17	2383 - 2450	68	II S6	16894	cag tgg ttg ctt ttc ctg ac	ttg cca gag aaa cat tct cc	130
18	2451 - 2557	107		16894	tga aca aag att cca cgt gc	ttc agg agc agg gt agc atc	170
19	2558 - 3367	810		16894	tag caa tgc tct aag tcc tg cgc agg aga acc gca aca a gc agc agg gag agc cgc agc	tgt ttc ctg agg aag tcc tc gcg atg acg tcg atg ctc tac cgt cat tct gcg gat tc	320 450 300
20	3368 - 3831	464		16894	ggt tct ttt tca ttc act tgc gag aat agc ctt atc gtc ac	ttt cct ggc agt ctt agc tg cag tga tgt gag agc aga g	430 200
21	3832 - 3973	142		16894 / 34275	:gg gaa att gtg gag gga gc	tga ctt ccg cca ccc tgg tg	250
22	3974 - 4103	130	III S1	16894 / 34275	agc ctg tgg tct gag tgg ac	tag gaa ggg gtg tgc tct gtg	210
23	4104 - 4163	60	III S2. III S3	16894 / 34275	atc cac tgc tct ctt gct tt	gtg gtt ctc act tat aat ctg c	170
24	4164 - 4270	107	III S3	34275	tgg cct cat tgg ctt ccc tgc	aag agg aaa ccc ttg cga ag	250
25	4271 - 4370	100	III S4	34275	cta ccc aac ctg acc tct gc	aca tga taa ccc tga cag tc	220
26	4371 - 4531	161	III S5	34275	ctc atg ctc tct gtc aac tc	tgg ttc aa tgg gaa tgt gc	250
27	4532 - 4669	138		34275	ctg ctt ccc aag cag tct ag	tcc tgg ata gat ttc cag tc	300
28	4670 - 4871	202	III S6	34275	agt ttt taa agg aca gat gg	ttt ccc tgc ccc att cct ttg c	280
29	4872 - 5036	165	IV S1	34275	ctc tgc cgc tct cac cac tg	ttt atc agg tag agg cag g	250
30	5037 - 5147	111	IV S1. IV S2	34275	ttc caa gcc cat agc tgt agc	tga ccc tgc tac tcc tgc ttc	180

31	5148 - 5231	84	IV S3	15496	act gtg cct cta aca tgc ac	aag tgc tgg ctc aag cag	250
32	5232 - 5348	117	IV S4	15496	tct gtg agt ggt gac agc tc	gtc acc tgt ctt ctc agc	240
33	5349 - 5414	66	IV S5	15496	tgg aag gac tct ggc acg tg	gga ggc tct ggg aac ctt ag	250
34	5415 - 5530	116		15496	aga agc cac tgg agg aat ggc	att atc aga gca ggt ccc ctt c	250
35	5531 - 5681	151	IV S6	15496	tcc gag tct ctg att tct cc	aga cgg ccc tca cag tgt c	210
36	5682 - 5809	128	IV S6	15496	ttc att ccc tcg gtc tct gc	ctg act gaa cct gtg aga c	350
37	5810 - 5906	97		15496	tgt gaa ccc att gcc tgc a	tgg gaa tga ctg cgc ttg c	200
38	5907 - 6012	106		15496	atg cct ggg aat gac tgc	tgt cac gcc tgt ctg tgc	200
39	6013 - 6120	108		15496	tga cac cca ggc agg cag	tct gtc ctg gtg gat tgg atc	200
40	6121 - 6221	101		15496	ttg gtg agc tca ccg tgt	ttc ccg tgg tga cat gca agc	200
41	6222 - 6331	110		15496	gtc cac aca ctg ctc tct gc	aca ctc cac ctc cct ggc	320
42	6332 - 6470	139		15496	gcc agg gag gtg gag tgt	ggt tcc ttc cac cgc aac	550
43	6471 - 6584	114		15496 / 30762	caa ctc ccc aat ggc tc	cct acc cag tgc aga gtg agg	350
44	6585 - 6620	36		15496 / 30762	tct gtg tgc acc atc cca tg	aag gat tgg gct cca tgg ag	200
45	6621 - 6807	187		15496 / 30762	gtt ggt gct agc tgc tga c	ctt tct tct tcc tta gtg tc	330
46	6808 - 7061	254		15496 / 30762	gtg tgc tgt ctg acc ctc ac	agc ctg ggg tca ctt gca gc	320
47	7062 - UTR	≥350		/ 30762	cct ttg ttt caa ttt tcg tgt ag	tgg ggc ctg ggt acc tcc aa	280

Note. Sizes of exons and PCR products are given in basepairs;
       domains of protein are indicated according to Stea et al., 1995.

Table 2 Polymorphisms in coding sequence of the invented gene

Location	Nucleotide change			Frequency	Consequence
exon 4	nt 854	G – A	Thr_{193}	0.02	-
exon 6	nt 1151	A – G	Glu_{292}	0.07	-
exon 8	nt 1457	G – A	Glu_{394}	0.38	-
exon 11	nt 1635	G – A	Ala_{454}	0.02	Ala_{454} – Thr (A454T)
exon 16	nt 2369	G – A	Thr_{698}	0.12	-
exon 19	nt 3029	G – A	Glu_{918}	0.07	-
exon 23	nt 4142	T – C	Phe_{1289}	0.22	-
exon 46	nt 6938	T – C	His_{2221}	0.46	-
exon 47	nt 7213	$(CAG)_n$	3'UTR	#	-

Note. Frequency as observed in 100 control chromosomes: # Seven alleles of $(CAG)_n$ were observed in the range between n=4 to n=14, with a heterozygosity value of 0.75.

Table 3 Mutations of the invented gene in families with FHM or EA-2

Disease	Family	Location	Domain	Nucleotide change		Consequence	
FHM	It-II	exon 4	I S4	nt 850	G - A	Arg_{192} - Gln (gain of SfcI site)	R192Q
FHM	US-P	exon 16	P-segment	nt 2272	C - T	Thr_{666} - Met	T666M
FHM	UK-B	exon 17	II S6	nt 2416	T - C	Val_{714} - Ala (gain of BpvI site)	V714A
FHM	NL-A/US-C	exon 36	IV S6	nt 5706	A - C	$Ile_{.61.}$ - Leu (gain of MnlI site)	I1811L
EA-2	CAN-191	exon 22	III S1	nt 4073	deletion C	frameshift (loss of NlaIV site)	$STOP_{1294}$
EA-2	CAN-26	intron 24	splice site	nt 4270-1	G - A	AC/gt - AC/at (loss of BsaAI site)	aberrant splicing

Table 4. Parental transmission of migraine for 36 unrelated Dutch families.

parents	N	offspring	N	affected N(%)	ratio*
heathy father x migraine mother	51	daughters	72	48 (66.7%)	2.3:1
		sons	72	21 (29.2%)	
migraine father x healthy mother	18	daughters	26	17 (65.4%)	2.5:1
		sons	15	4 (26.7%)	

* ratio of proportion affected sons/proportion affected daughters

32

CLAIMS

1.   An isolated and/or recombinant nucleic acid encoding a
Ca^{2+}-channel α1 subunit related to (familial hemiplegic)
migraine and/or episodic ataxia type-2 derived from, related
to or associated with a gene which in humans is present on
5   chromosome 19p13.1-19p13.2 or a specific fragment or homolog
or derivative thereof.

2.   A nucleic acid according to claim 1 which is a cDNA
molecule.

3.   A cDNA molecule according to claim 2 comprising a
10   6800 bp coding region.

4.   A nucleic acid according to claim 1, 2 or 3 which is of
human origin.

5. . A nucleic acid according to claim 4 and showing at least
70% homology with the nucleic acid sequence as listed in
15   figure 1.

6.   A nucleic acid according to any of claims 1-5 and
showing at least 90% homology with the nucleic acid sequence
as listed in figure 1.

7.   A nucleic acid according to any of claims 1-6 and
20   showing a mutation at codon 192 resulting in the replacement
of arginine by glutamine.

8.   A nucleic acid according to any of claims 1-7 and
showing a mutation at codon 666 resulting in the replacement
of threonine by methionine.

25   9.   A nucleic acid according to any of claims 1-8 and
showing a mutation at codon 714 resulting in the replacement
of valine by alanine.

10.   A nucleic acid according to any of claims 1-9 and
showing a mutation at codon 1811 resulting in the replacement
30   of isoleucine by leucine.

11.   A nucleic acid according to any of claims 1-10 and
comprising a CA-repeat sequence as shown in figure 2.

12. A nucleic acid according to any of claims 1-11 and comprising a (CAG)n repeat sequence as shown in figure 2.

13. A nucleic acid according to any of claims 1-12 and comprising a polymorphism in the coding sequence.

14. A nucleic acid according to claim 13 and comprising a polymorphism in the coding sequence as shown in table 2.

15. A nucleic acid according to claim 13 or 14 and comprising a mutation at codon 454 resulting in a replacement of alanine by threonine.

16. A nucleic acid according to any of claims 1-15 and comprising a deletion.

17. A nucleic acid according to any of claims 1-16 and comprising a frameshift at codon 1266.

18. A nucleic acid according to any of claims 1-17 and comprising a mutation resulting in abberant splicing.

19. A nucleic acid according to any of claims 1-18 and comprising a mutation resulting in abberant splicing of intron 28.

20. An isolated and/or recombinant nucleic acid encoding a CA^{2+} channel subunit or a functional fragment thereof related to (familial hemiplegic) migraine and/or episodic ataxia type 2.

21. An isolated and/or recombinant nucleic acid encoding a CA^{2+} channel β2 subunit related to (familial hemiplegic) migraine and/or episodic ataxia type 2, derived from, related to or associated with a gene which in humans is present on chromosome 10p12 or a specific fragment thereof.

22. A method for localising or identifying a gene using a nucleic acid molecule or a fragment of fragments thereof according to any of claims 1-21.

23. A method according to claim 22 wherein the gene is related to episodic neurological disorders.

24. A method according to claim 22 or 23 wherein the gene is related to migraine.

25. A method according to claim 22, 23 or 24 wherein the gene is related to FHM and/or EA-2 and/or autosomal dominant cerebellar ataxia.

34

26. A method distinguishing between alleles of a gene using a nucleic acid molecule or a fragment of fragments thereof according to any of claims 1-21.

27. A method according to any of claims 23-26 in which the

5   gene is of human origin.

28. A method according to any of claims 23-27 which comprises selecting a cell or an animal.

29. A recombinant expression vector comprising a nucleic acid molecule or a fragment of fragments thereof according to

10 any of claims 1-21.

30. A cell or an animal comprising a vector according to claim 29.

31. A cell or an animal comprising a nucleic acid molecule or a fragment of fragments thereof according to any of claims

15 1-21.

32. A cell or an animal selected by a method according to claim 28.

33. A cell or an animal comprising a genome in which nucleic acid sequences corresponding to nucleic acid molecules

20 according to any of claims 1-21 have been modified.

34. Use of a cell or an animal according to any of claim 30-33 to test or develop specific medication for the treatment of FHM, EA-2, SCA6, migraine or other neurological disorders associated with cation channel dysfunction.

25 35. A protein or peptide comprising an amino acid sequence encoded by a nucleic acid molecule, or a fragment or fragments thereof, according to any of claims 1-21.

36. A natural or synthetic antibody directed against a protein or peptide according to claim 35.

30 37. Use of a protein or peptide or antibody according to claim 35 or 36 to detect or diagnose FHM, EA-2, SCA6, migraine or other neurological disorders associated with cation channel dysfunction.

Figure 1

```
//
Submission no   :   1
exon 1          :   <..672
start codon     :   381..383
intron 1        :   673.
Remarks         :   no consensus splice site intron 1
```

```
tttttttacg ttctcttttt tttcgagtgg tgactggatg ctgattcttc        50
ctcgtatttt tgctgcttct ctctccctcc cctccttccc gggcccgggc       100
ccgccccgca ccctccttcc gccctcctt ctccggggtc agccaggaag        150
atgtcccgag ctgctatccc cggctcggcc cgggcagccg ccttctgagc       200
ccccgacccg agcgccgagc cgccgcgcga tgggctgggc cgtggagcgt       250
ctccgcagtc gtagctccag ccgccgcgct cccagccccg gcagcctcag       300
catcagcggc ggcggcggcg gcggcggcgt cttccgcatc gttcgccgca       350
gcgtaaccgg agccctttgc tctttgcaga ATGGCCCGCT TCGGAGACGA       400
GATGCCGGCC CGCTACGGGG GAGGAGGCTC CGGGGCAGCC GCCGGGGTGG       450
TCGTGGGCAG CGGAGGCGGG CGAGGAGCCG GGGGCAGCCG GCAGGGCGGG       500
CAGCCCGGGG CGCAAAGGAT GTACAAGCAG TCAATGGCGC AGAGAGCGCG       550
GACCATGGCA CTCTACAACC CCATCCCCGT CCGACAGAAC TGCCTCACGG       600
TTAACCGGTC TCTCTTCCTC TTCAGCGAAG ACAACGTGGT GAGAAAATAC       650
GCCAAAAGAT CACCGAATGG CCatatcctt ttgcccgaac cccagcagca       700
gctgcgcctc cccctcctcc ctccgcctcc cctcttccag gctgggagag       750
agacccgggg gttgatggga ggtggggagg aggggggtct tccagggget       800
gggagagggg gcaccgggag gagtgtgaaa gaatctctcc accccgagct       850
gggttgagct accctggagg cttgggaatg ggttttcgg gggctggggg        900
ccggccagcc ggagagtgga tccttcccaa ggaccgactc tagaatgaga       950
tct                                                          953
```

```
//
Submission no      :      2
Intron 1           :    <..88
Exon 2             :    89..194
Intron 2           :    195..>
Remarks            :    No consensus splice site intron 1

        gatctttycc actggggtca gtgggggtgg gtgcacctcc aacacccttc        50
        ttttctttga acaagatttt tccttaattc cccaatacTC CCTTTGAATA       100
        TATGATTTTA GCCACCATCA TAGCGAATTG CATCGTCCTC GCACTGGAGC       150
        AGCATCTGCC TGATGATGAC AAGACCCCGA TGTCTGAACG GCTGgtgagt       200
        gatgtctttt ctcagggtct tctccttggc tttagcagga cattaatttt       250
        tgggggagtg gagcagggca cagaggaggc tctcagtcct ggagcccaga       300
        gccagatcat gggaagccta aatttccttt tcattttttc ttgaaccaga       350
        gtctcgctct gtcacccagg ctggagtgca gtggttcagt catagctcac       400
        tgcagcctcc acctcctggg ctcaagccat cctcccactg cagcctcctg       450
        agtagcaggg actaacaggt gccaccatgc ccagttaatt ttcttatttt       500
        tatctttttt tgtaagaaga tggggat                               527
```

```
//
Submission no  :    3
Intron 2       :    <..57
Exon 3         :    58..197
Intron 3       :    198..>

    gatcttgtca  acatctgccc  agcccaagac  gctgaccttg  ccttctctcc        50
    cttccagGAT  GACACAGAAC  CATACTTCAT  TGGAATTTTT  TGTTTCGAGG       100
    CTGGAATTAA  AATCATTGCC  CTTGGGTTTG  CCTTCCACAA  AGGCTCCTAC       150
    TTGAGGAATG  GCTGGAATGT  CATGGACTTT  GTGGTGGTGC  TAACGGGgta       200
    agtggcgcgt  gctatacgct  ttggatttaa  ctagctgaag  gattacgagg       250
    cttttggttg  gtgtggtccg  ggccaggctc  aggaaggctg  agcccttgtg       300
    ttctccctcc  ccttgttatg  cgcctgcctc  ctttctgcca  acaccccacc       350
    tccatgtctc  agctgtatat  tacagcagat  gctttctgtt  acaattaaaa       400
    taatagctca  ttattgttgg  ctgcttccag  agtgctttat  g                441
```

```
//
Submission no  :    4
Intron 3       :    <..142
Exon 4         :    143..234
Intron 5       :    235..>
```

```
       aaaactgagg  ccagtggtgt  cgagtcacct  gcctgtggtc  acccaaccaa      50
       tacaggacag  cttggaatcc  caagcacccc  cgccctgctg  tctgaccccc     100
       aaaacccacc  ctctgttctc  cattctggct  tctttctttc  agCATCTTGG     150
       CGACAGTTGG  GACGGAGTTT  GACCTACGGA  CGCTGAGGGC  AGTTCGAGTG     200
       CTGCGGCCGC  TCAAGCTGGT  GTCTGGAATC  CCAAgtgcgt  gagtttccga     250
       ccctgacaa                                                      259
```

```
//
Submission no    :    5
Intron 4         :    <..118
Exon 5           :    119..271
Intron 5         :    272..>


    cttaatattc cctcaggaac acacctgctt tgtctgggag agacctgggc        50
    gtcttggtgg cggggttttg ggggtacttg ctcatgggct tatggggcct       100
    ctctctgtgt cccccagGT TTACAAGTCG TCCTGAAGTC GATCATGAAG         150
    GCGATGATCC CTTTGCTGCA GATCGGCCTC CTCCTATTTT TTGCAATCCT        200
    TATTTTTGCA ATCATAGGGT TAGAATTTTA TATGGGAAAA TTTCATACCA        250
    CCTGCTTTGA AGAGGGGACA Ggtaggtcca cggagcatga tgcatctttc        300
    cagttttctc cttcagggac aagctcttgg gaggattagg caggggtgtg        350
    cttctttctc ctggcagctg ggaggaccgt ctccttcaga gagcactac         399
```

```
//
Submission no   :       6
Intron 5        :       <..22
Exon 6          :       23..216
Intron 6        :       217..>


        tttttttccct tccctttttgt agATGACATT CAGGGTGAGT CTCCGGCTCC       50
        ATGTGGGACA GAAGAGCCCG CCCGCACCTG CCCCAATGGG ACCAAATGTC      100
        AGCCCTACTG GGAAGGGCCC AACAACGGGA TCACTCAGTT CGACAACATC      150
        CTGTTTGCAG TGCTGACTGT TTTCCAGTGC ATAACCATGG AAGGGTGGAC      200
        TGATCTCCTC TACAATgtaa gtgatgctgg gacagtgtgt gtggacaatc      250
        agagtctcag ggaggtggcc tcctgggacc agtgagactc caaggctgca      300
        atggagggac cctgagctgg gaaaggcagc ccaaggacaa cacagcccca      350
        ctgaagctgg cctgaggctc aggcttttga agattacagg ggctcatgag      400
        cagaactcta actatagggc atagaagtct ggagggcccc cagatgcaac      450
        atcatttttc attgtgcaag tgtttagata taattttaga ttttgaata       500
        cggaaaggtt atgtgatcca aaatccaaca cagataaaag atagagtaat      550
        atctttggac gtaggcgagg ggtccctgcc ctgagg                    586
```

```
//
Submission no  :    7
Intron 6       :    <..183
Exon 7         :    184..287
Intron 7       :    288..>


     tttcttcaga aaacggttcc ttcctccatt tccccctctg ggatgccaga          50
     gccccagaac tccacaagcc aagaacattt aagacagagc cacaagagaa         100
     ccgagcttcc ccttccctca cctgtcaggt tctatctgag tcccagtcaa         150
     ctctcacctg ctttccctcc tcacacccta cagAGCAACG ATGCCTCAGG         200
     GAACACTTGG AACTGGTTGT ACTTCATCCC CCTCATCATC ATCGGCTCCT         250
     TTTTTATGCT GAACCTTGTG CTGGGTGTGC TGTCAGGgta agtttctgct         300
     actccccacc ccatcccact cactcctctt tgctaacttc tttccaagta         350
     gaggccattg aagctttgtt ttcattcact agacaga                      387
```

```
//
Submission no  :    8
Intron 7       :    <..190
Exon 8         :    191..306
Intron 8       :    307..>
Sequence       :    412
Remark         :    intron 7 contains CA-repeat (D19S1150)


   cccagtcttt tcccagaagt cctgactcct cctgttgaaa actcctgacc      50
   tccagggact tctgaatccc caaacacaca cacacacaaa cacacacaca     100
   cacacacaca cacacacaca caaacacaca cacaaacgtt tcctaacatt     150
   ttaaaacag ccatactctg gcttttctat gcttctccag GGAGTTTGCC      200
   AAAGAAAGGG AACGGGTGGA GAACCGGCGG GCTTTTCTGA AGCTGAGGCG     250
   GCAACAACAG ATTGAACGTG AGCTCAATGG GTACATGGAA TGGATCTCAA     300
   AAGCAGgtga ggccctttca tcctggggcc cagggatgga gatcccaggc     350
   cacagagtac aaagagagtc atgcagtttg gagaaggcta agctgggagg     400
   gttatgatgg ga                                             412
```

SUBSTITUTE SHEET (RULE 26)

```
//
Submission no    :    9
Intron 8         :    <..513
Exon 9           :    514..570
Intron 9         :    571..>
```

```
gagtaggaag ttagaggcag ggtggtcagg gaaggcttct ctaaggaagt    50
accctctgag cagagagacc tgaaggacgt gaagaaggaa gctgtgggga   100
tgtcaaggga aggggcattc caggcagaga cagcaagtgc aaaggccctg   150
agctaggaac gtatttgaga cacagcaagg aagccagtgc agctgaaaca   200
gagtgagagg tggggacagc tggaggagag gaagacagga aggtgatgga   250
gatcagatca agcaggggct tataggctgt ggtgtggaca ttggtttttta  300
ttttgcgcga ggtggggaga atgttggcta ttgctactgt tgcggaggtg   350
gggcttgaag tcacaaacca cccagcagca tgttttttgg tcggttgagc   400
tgtcaccatc agtcagcaga gaatgggggt ggccgggcag acccttcttc   450
ctggtccaag ggagaactca tcctccaaat gcaggagctt aactctgtgc   500
tcttcctctt cagAAGAGGT GATCCTCGCC GAGGATGAAA CTGACGGGGA   550
GCAGAGGCAT CCCTTTGATG gtaactgctc taaacccacc tcaggggtgg   600
gtcccagggg a                                            611
```

```
//
Submission no    :    10
Intron 9         :    <..86
Exon 10          :    87..179
Intron 10        :    179..>
```

```
ttaatccaag acacactgtg tgtcctatat ggtctgtgtt cgaaaaaggg      50
taacgtcttt ttctcttgcc atgtttccat tgttagGAGC TCTGCGGAGA     100
ACCACCATAA AGAAAAGCAA GACAGATTTG CTCAACCCCG AAGAGGCTGA     150
GGATCAGCTG GCTGATATAG CCTCTGTGGg tgagtccctt cctctgccac     200
ctatcagttg ttcatcacct atcgcccaag agacatggtg gggtggggggc    250
agagggcttg caaaccgtgc tgcctggatt tgggtctcag ctccacccctt    300
tcccacctgt gcgtgtgtcc tgggcagatt acatcattat gggaataaca    350
tccgtgccta gcttctcatt attttgtggg aattcaacta aatgatcccc     400
atgaagcatg gcaaaccagc acctggcagg gacgaagctc ccagtcaagt     450
tggtgaatgt ttgtgactca ttcgggaagt atcatggggg acctgcttat     500
attaggtgct tggttgcaaa caaacaaggc agtcacgagg ctgagctggg     550
aggatcactt gagcctggga agtggaggct gcaataagcc attattgtgt     600
tactgcactc cagcctggca cagaaaaaaa aaaaaaanac aaactgagcc     650
agcaca                                                    656
```

```
//
Submission no   :    11
Intron 10       :    <..450
Exon 11         :    451..660
Intron 11       :    661..>


        gatcacttct aaagttaaat gtccatggga aaacagtctc atccacatct        50
        ctttctggag gccttccaag cgtgctccat gcagctctgt tgcctgcccc        100
        tgcatcaggg aatggaggct ctgctttatc ctgccctgtg gtgtgactcc        150
        cagaggcatc agatgtggct gggagtggga gacatggaaa attggctcct        200
        gcaacagaga actatcagcc ttcccatcaa ttggttactt ctaattctgt        250
        tatttttcag gggcactgtc ttctcataag ctccatctat gcaaaactaa        300
        gcccatgggt catgatggtt ccctcaggcc agaggcttgc tggagagact        350
        aatggatccc ctggctaaaa tctgtgcttg ggctgcacat tggttaattt        400
        cttctgaagg aacagcctga gcctgacatt ctccatcttt tccctggcag        450
        GTTCTCCCTT CGCCCGAGCC AGCATTAAAA GTGCCAAGCT GGAGAACTCG        500
        ACCTTTTTTC ACAAAAAGGA GAGGAGGATG CGTTTCTACA TCCGCCGCAT        550
        GGTCAAAACT CAGGCCTTCT ACTGGACTGT ACTCAGTTTG GTAGCTCTCA        600
        ACACGCTGTG TGTTGCTATT GTTCACTACA ACCAGCCCGA GTGGCTCTCC        650
        GACTTCCTTT gtgagtatca cccagcccca cccctgccaa ctccctgatc        700
        cctccctcac accctttttc cacttctctt tctctggtag tatgtgtatc        750
        ttctttggtc ctcattgaat ctgccctt                                 778
```

```
//
Submission no   :     12
Intron 11       :     <..323
Exon 12         :     324..436
Intron 12       :     437..>


        gatcacttgt ggccaggagt tcaagancag ccagggcaac atagtgagga        50
        cccccatctc cacattaaaa attttaaaaa gaaaaaagat aagtcagaag       100
        ttgggtgtgg tgacacatgc ctgtagttct agcatgttgg aggccaaatc       150
        agggaaactg tttgaggcca ggagtttgaa accagcctaa cagcatagca  .    200
        agacctcatc tctacaaaaa ataaaaagtt taaaaatgat aataaaagga       250
        aagtcagagc cacctggaac ccctaccctc agcaagccta acctcctctc       300
        tgtttcctcc ttctcccttc tagACTATGC AGAATTCATT TTCTTAGGAC       350
        TCTTTATGTC CGAAATGTTT ATAAAAATGT ACGGGCTTGG GACGCGGCCT       400
        TACTTCCACT CTTCCTTCAA CTGCTTTGAC TGTGGGgtaa gtgctcttgt       450
        ttctaagagt tcatttctcc agctcttgcc tggaatgaca gatacctgga       500
        cacattaaag ggagaaaggt aaagtcaccc ctgaatatga gagactcaga       550
        tggatgcaga aggaatgaga aaacaatcca aacactggca aggatacagt       600
        gtacccagaa ccctcaacca ccgcca                                 626
```

```
//
Submission no    :    13
Intron 12        :    <..545
Exon 13          :    546..658
Intron 13        :    659..821
Exon 14          :    822..953
Intron 14        :    954..>
```

```
        gatcngncat gcacaccagc ctgggtgata agagcaagac tcctctcaaa          50
        ataaatgaat aaataaaaat aaataaataa ataagaggcc gggtgcagtg         100
        gctcaatgct ttggaaagtg gaggccaaca gttggagaga ccaaagcagg         150
        aggatggctt cagcccagaa gtttgaggcc mgcctgggca atactagcga         200
        gacactatct ctataaaaat gttttaaaat tagccagatg tggtggggca         250
        cacctgtaat cccagctact caagaggctg aggtgggagg atcacttaag         300
        cccaggagga cagtgctgca gtgagctatg attgcgccca ctgcactcca         350
        gcctgggtga cacagtgaga cccggtctct atagataaat gaatggatga         400
        atgagggggt caaggatcct cacccggctt ccatttggag ggaggagttt         450
        ggttgagttc ttgcaaggtt ggtacctagg aaatgcttgc cagttctgga         500
        gcccagacac tgtccctgga catgagacca ggttctctgc cctagGTTAT         550
        CATTGGGAGC ATCTTCGAGG TCATCTGGGC TGTCATAAAA CCTGGCACAT         600
        CCTTTGGAAT CAGCGTGTTA CGAGCCCTCA GGTTATTGCG TATTTTCAAA         650
        GTCACAAAgt aagtctttgg ggttcctgga catttgtaca gggggtgggg         700
        atgggggaca tggtgggggc gcctccagaa agttgggaaa gtgagcctcg         750
        tgtttcgagg gctgactccg gggcctgcct wccccgcctg gcctgagtcc         800
        tcgcctggsc tctgtcggca gGTACTGGGC ATCTCTCAGA AACCTGGTCG         850
        TCTCTCTCCT CAACTCCATG AAGTCCATCA TCAGCCTGTT GTTTCTCCTT         900
        TTCCTGTTCA TTGTCGTCTT CGCCCTTTTG GGAATGCAAC TCTTCGGCGG         950
        CCAgtaagtc cttcacagga attcaa                                    976
```

```
//
Submission no   :    14
Intron 14       :    <..201
Exon 15         :    202..274
Intron 15       :    274..>
```

```
        ccctccacgt gcaggctgcc ttcctcgtag cccagacacc catttgcggt        50
        cacccaaatg ggcagggccc tgggtaccac tcagggtttc ctggggacag       100
        agatgatgga aacgttcgtt tccttggaga tgagatactg agccacaccc       150
        tcagagcacc ccgggtgggg ccaacgtgaa atgtctgtgt cctccctgca       200
        gGTTTAATTT CGATGAAGGG ACTCCTCCCA CCAACTTCGA TACTTTTCCA       250
        GCAGCAATAA TGACGGTGTT TCAGgtacag cctccacctg gccccacggg       300
        ccaacacctc tcagtgtcac agatgaaagt gcctgctcca catccaaggg       350
        gcttccctga actcctcctt ctctacctgg cctttcaca  ccactttgaa       400
        acacagattt tatggttatc attattcaat tatggtgagg ccaacagatc       450
        aggagatgaa tgtcattgga aagatagttt gtggctgggc acggtggctc       500
        acacccataa tcccagcact ttggccaggt acggtggctc acacctgtaa       550
        tcccaacgct ttgggaagcc cagtgggcgg atcacttga  gatcaggaat       600
        tcgagaccag cctggccaan atggtgaaac cccatctcta ctaaaaatac       650
        aaaaattagc cgggcgtggt agcacatgcc tgtaatccca gctactcggg       700
        agatgaggca caagaattgc ttgaacctgg gaggcagagg ttgcagtgag       750
        ccaagatcgc gccactgcac tcmagcctgg caacagagt  gagactccat       800
        ctcaaaaaag caaaagaaaa aaaaaccac  tttgggaggt caagatggga       850
        ggactacttg aggccaggag tttgagacaa gtctgggcaa catagtgaga       900
        ctccgtctct gcaaaaaaat wataataata attagctggg catggtgata       950
        catacctcct agctactagg gcagctgaag tggaaggatt gcttaagccc      1000
        aggaggttga ggctgcagta agctacaatc acaccactat actccagcct      1050
        gggcgagaga gcaaagccct gtctcaaaaa cgaaaagaaa gtttgttata      1100
        ctcacagatc                                                  1110
```

```
//
Submission no   :    15
Intron 15       :    <..524
Exon 16         :    525..642
Intron 16       :    643..795
Exon 17         :    796..863
Intron 17       :    864.>
```

```
        gatcctccca cccttggcctc ccaaagtgct gggattacag gcatgagcca        50
        tggcatgcgg tctcttcctg ttcttataag ggcactaata ccatcatgaa       100
        gtcccccatg acctcatcta accctagtta cctcttaaag gccccatctc       150
        caaataccat cccatcatag gttagggctt caactcatga atttggaggc       200
        gggcacaatt tagtccataa caaatcccct taatcacatc aagtaagaca       250
        gagttacagg agggtctgtg actcctccag ggtcccattt tcctagaagc       300
        caggctaaga gccccacgac gcaggaacgg ccctttctac tcgcaaacaa       350
        agagaaaagc caaggagaag ccaacacgga gtctggctct gcaaaccggg       400
        caggattgtt aaagacctcc tgggctcggg gatggggtgg gcggattccg       450
        gctccacagc tgcatctcca aggggcccgt ggctgagagg ggggttggct       500
        gtgtgtttct tcctcccctt tcagATCCTG ACGGGCGAAG ACTGGAACGA       550
        GGTCATGTAC GACGGGATCA AGTCTCAGGG GGGCGTGCAG GGCGGCATGG       600
        TGTTCTCCAT CTATTTCATT GTACTGACGC TCTTTGGGAA CTgtatcctt       650
        catggagaga gagaagggga caggcctgga cctctggcag aggagaggtt       700
        gcagggctc aagggagggt actgagagac ccagataccc agggcccaag       750
        tggtgtccca ccagtggttg cttttcctga ctcagacatt tgcagACACC       800
        CTCCTGAATG TGTTCTTGGC CATCGCTGTG GACAATCTGG CCAACGCCCA       850
        GGAGCTCACC AAGgtggagg cggtgggaga atgtttctct ggcaaagtta       900
        ccacctgccc atggcagatc aagcactttt ttggattaac tgagccacag       950
        gaaataacat tttcaaatag atkaaaaaga tc                          982
```

```
//
Submission no   :   16
Intron 17       :   <..119
Exon 18         :   120..226
Intron 18       :   227..>
```

```
          ccttggttct gattggtcga aatatttcaa atgttgcccc tggtcagcaa            50
          cagggtcaga agtgagtccc caaggcctag ttcatgtttt gtgaacaaag           100
          attccacgtg cctttttcagG ACGAGCAAGA GGAAGAAGAA GCAGCGAACC          150
          AGAAACTTGC CCTACAGAAA GCCAAGGAGG TGGCAGAAGT GAGTCCTCTG           200
          TCCGCGGCCA ACATGTCTAT AGCTGTgtaa gtcccctaat ccctgggatg           250
          cta⌐cctggc tcctgaacgt gtccgaccac tatccaggca cagatttggg           300
          aagcagtggg ggtg                                                  314
```

```
//
Submission no  :    17
Intron 18      :    <..209
Exon 19        :    210..1019
Intron 19      :    1020..>
```

```
        gcccctagcc aggtgggagc catggagggt tcttgagcag aggaggctgg        50
        gacctgactc agatgctcac agactcctag cattcaggtg gggagtagag       100
        ggtggagagc aggagtggga ggctgagatg tgggttggtt cgcctgggtc       150
        atccatccaa gctacagtgc ctagcaatgc tctaagctcc tgtgaccatg       200
        ccactgcagG AAAGAGCAAC AGAAGAATCA AAAGCCAGCC AAGTCCGTGT       250
        GGGAGCAGCG GACCAGTGAG ATGCGAAAGC AGAACTTGCT GGCCAGCCGG       300
        GAGGCCCTGT ATAACGAAAT GGACCCGGAC GAGCGCTGGA AGGCTGCCTA       350
        CACGCGGCAC CTGCGGCCAG ACATGAAGAC GCACTTGGAC CGGCCGCTGG       400
        TGGTGGACCC GCAGGAGAAC CGCAACAACA ACACCAACAA GAGCCGGGCG       450
        GCCGAGCCCA CCGTGGACCA GCGCCTCGGC CAGCAGCGCG CCGAGGACTT       500
        CCTCAGGAAA CAGGCCCGCT ACCACGATCG GGCCCGGGAC CCCAGCGGCT       550
        CGGCGGGCCT GGACGCACGG AGGCCCTGGG CGGGAAGCCA GGAGGCCGAG       600
        CTGAGCCGGG AGGACCCCTA CGGCCGCGAG TCGGACCACC ACGCCCGGGA       650
        GGGCAGCCTG GAGCAACCCG GGTTCTGGGA GGGCGAGGCC GAGCGAGGCA       700
        AGGCCGGGGA CCCCCACCGG AGGCACGTGC ACCGGCAGGG GGGCAGCAGG       750
        GAGAGCCGCA GCGGGTCCCC GCGCACGGGC GCGGACGGGG AGCATCGACG       800
        TCATCGCGCG CACCGCAGGC CCGGGGAGGA GGGTCCGGAG GACAAGGCGG       850
        AGCGGAGGGC GCGGCACCGC GAGGGCAGCC GGCCGGCCCG GGGCGGCGAG       900
        GGCGAGGGCG AGGGTCCCGA CGGGGGCGAG CGCAGGAGAA GGCACCGGCA       950
        TGGCGCTCCA GCCACGTACG AGGGGGACGC GCGGAGGGAG GACAAGGAGC      1000
        GGAGGCATCG GAGGAGGAAg taagtggagg tgacctcgaa tccgcagaat      1050
        gacggtaaca ttaataatac aacagccaaa gtagcacgtg ctgtgtattt      1100
        gttataaaat ata                                            1113
```

//
Submission no : 18
Intron 19   : <..67
Exon 20   : 68..531
Intron 20   : 532..>

```
       gtcctgaaac  tttgcctttt  aatcctaaat  cattgttggt  tctttttcat          50
       tcacttgcct  tcctcagAGA  GAACCAGGGC  TCCGGGGTCC  CTGTGTCGGG         100
       CCCCAACCTG  TCAACCACCC  GGCCAATCCA  GCAGGACCTG  GGCCGCCAAG         150
       ACCCACCCCT  GGCAGAGGAT  ATTGACAACA  TGAAGAACAA  CAAGCTGGCC         200
       ACCGCGGAGT  CGGCCGCTCC  CCACGGCAGC  CTTGGCCACG  CCGGCCTGCC         250
       CCAGAGCCCA  GCCAAGATGG  GAAACAGCAC  CGACCCCGGC  CCCATGCTGG         300
       CCATCCCTGC  CATGGCCACC  AACCCCAGA   ACGCCGCCAG  CCGCCGGACG         350
       CCCAACAACC  CGGGGAACCC  ATCCAATCCC  GGCCCCCCCA  AGACCCCCGA         400
       GAATAGCCTT  ATCGTCACCA  ACCCCAGCGG  CACCCAGACC  AATTCAGCTA         450
       AGACTGCCAG  GAAACCCGAC  CACACCACAG  TGGACATCCC  CCCAGCCTGC         500
       CCACCCCCCC  TCAACCACAC  CGTCGTACAA  Ggtgagaccc  tctgctctca         550
       catcactggg  cagggacct   ggcgtcctgg  agccagaggt                     590
```

```
//
Submission no   :   19
Intron 20       :   <..75
Exon 21         :   76..217
Intron 21       :   218..>
```

```
          ggagtacacc gaggagttcc cagagacttg tgggaaattg tggagggagc        50
          cctgtgttgg ttcttgtccc aacagTGAAC AAAAACGCCA ACCCAGACCC       100
          ACTGCCAAAA AAAGAGGAAG AGAAGAAGGA GGAGGAGGAA GAAGACGACC       150
          GTGGGGAAGA CGGCCCTAAG CCAATGCCTC CCTATAGCTC CATGTTCATC       200
          CTGTCCACGA CCAACCCgtg agtatggccc ccgagcagag ggcaggggggg      250
          gctgggtctc ccaccagggt ggcggaannn nnnnnnnnn nnnnnnnctc        300
          ccaccagggt ggcggaagtc aggccagatt agagggcaat                 340
```

```
//
Submission no   :   20
Intron 21       :   <..97
Exon 22         :   98..227
Intron 22       :   228..>


    gatctcagta  gtggtaggta  acatgagatt  atggaagaaa  agggtttgtg        50
    agcctgtggt  ctgagtggac  ctctgcacgc  ccatctgtct  ccaacagCCT       100
    TCGCCGCCTG  TGCCATTACA  TCCTGAACCT  GCGCTACTTT  GAGATGTGCA       150
    TCCTCATGGT  CATTGCCATG  AGCAGCATCG  CCCTGGCCGC  CGAGGACCCT       200
    GTGCAGCCCA  ACGCACCTCG  GAACAACgtg  agtcccacag  agcacacccc       250
    ttcctagcct  ggctgctctg  cctcaggcca  ctttctcctg  catccaaaat       300
    gctcataggt  agggtgggat  gttggggtca  cccctaggca  tagcccttat       350
    ggctgctggt  tgagagggga  agctctgatt  ccttggggat  gctcttggga       400
    gcaagacatt  ccttgaggca  gtttctctgt  gagcctggtg  gggtggaggt       450
    ggcccagagt  gactggggct  gaaaatt                                  477
```

```
//
Submission no    :     21
Intron 22        :     <..33
Exon 23          :     34..93
Intron 23        :     94..>


     gatccactgc tctcttgctt ttatccctta cagGTGCTGC GATACTTTGA          50
     CTACGTTTTT ACAGGCGTCT TTACCTTTGA GATGGTGATC AAGgtgagtg          100
     cagattataa gtgagaacac acggtaattt tttttttttaa gcaagtgcag    ·    150
     ggctgggcac agtggatc                                            168
```

```
//
Submission no    :    22
Intron 23        :    <..232
Exon 24          :    233..339
Intron 24        :    340..>


      gatctaagag  ccggcaagcc  agagctggct  tccatcaggc  aaaggggggc       50
      cgcctcatgg  ggcaggggct  ccccactcct  ccctgggagt  cctctggcca      100
      ctgcccatcc  ctgcaagatg  aggtggcctc  attggcttcc  ctgcctctcc      150
      ccgagaggct  agagagtggg  tggcagcacc  ccagggtggg  gatcaggtgg      200
      gggttctgag  caccctctct  tctcccccac  agATGATTGA  CCTGGGGCTC      250
      GTCCTGCATC  AGGGTGCCTA  CTTCCGTGAC  CTCTGGAATA  TTCTCGACTT      300
      CATAGTGGTC  AGTGGGGCCC  TGGTAGCCTT  TGCCTTCACg  taagtctctt      350
      cgcaagggtt  tcctcttg                                            368
```

```
//
Submission no    :    23
Intron 25        :    <..244
Exon 25          :    245..344
Intron 24        :    345..>
```

```
gatcttaacc ccaagacact tcatctaaag gaaaaactgc cataatacac        50
agattatttt aggtcagctc actttactgc catctgctgg gaagttgtaa       100
taatacaaat atccatacac gatggctagg atgttatcag cacctccttt       150
aatgtgttgt ccttgagcag tgtacaacct gctcagctgt acatgataac       200
cctgacagtc ccccccaccg caccccacca tctcccaatc tcacCTTGAG       250
CTTTGGCAGC CGCTTGATGG TTTTAAGAGG TCGTAGCACC CGGAGGACTC       300
GGAGGGATTT AATCGTGTTG ATGTCTTTTC CTTTGCTATT GCCActgtgg       350
aggaatgttt aggtgggaag aagggaagag aggaagcaga ggtcaggttg       400
ggtaggggge agcccacagc tccatgggac cctacccttc ccaggcctag       450
aagtctgggg tgagcttggc acaagcctgc cctttcctgg tgaagagtgg       500
tccattttac cctgt                                            515
```

```
//
Submission no    :    24
Intron 25        :    <..67
Exon 26          :    68..228
Intron 26        :    229..>
```

```
        ggccactgga ggcagaaggt tggcaggtcc ccagcccctc atgctctctg          50
        tcaactccac cccacagGCT GTGTTTGACT GTGTGGTGAA CTCACTTAAA         100
        AACGTCTTCA ACATCCTCAT CGTCTACATG CTATTCATGT TCATCTTCGC         150
        CGTGGTGGCT GTGCAGCTCT TCAAGGGGAA ATTCTTCCAC TGCACTGACG         200
        AGTCCAAAGA GTTTGAGAAA GATTGTCGgt gggtctccgc ttccagcac          250
        attcccattg gaaccagcag gtgggcaggg gggaagtggc tagaggcatt         300
        ggccacttgg gctcagagac tggagaagtg atgagccttg gaagtgactc         350
        agttgcaacc agcttggatc aagggtagaa agaaaaccgg ttttagaatt         400
        tgagtc                                                          406
```

```
//
Submission no   :    25
Intron 27       :    <..177
Exon 27         :    178..315
Intron 26       :    316..>
Remark          :    reversed direction!

        gatctcaaac tcctggcctc aagtgataca tctgccttgg cctcctaaag        50
        tgttgggatt acaggcgtga gcaccatgcc cggcctccaa gacctttctt       100
        attgctaagc tctcaggccc tttatcctcc tgctccccag ggctcctcct       150
        ggatagattt ccagtcgggc cacttacTGT GGCCAGCCTT CTCCCGTGGA       200
        CACGGTGAAG AGGGTCAGCA GAGCCCACAG CACATTGTCG TAATGGAATT       250
        CATACTTCTT CCACTCCCGG TCTCGCGCCT TCACCTCATT CTTCTCGTAG       300
        AGGAGGTATT TGCCTctgcc acagagagtg gggactgtta gtaaatggga       350
        aagagggct gtcttgcact tgtctttggt tatcagagac aggggggaggg       400
        aaaggaagag tggtccacca ncctagactg cttgggaagc agtgacttcc       450
        catcctgcca ccatgtgttc ctgtgcttca tagggatgn cgtgtgcaat        500
        ctactttna ggataa                                            516
```

//
Submission no   :   26
Intron 27   :   <..84
Exon 28   :   85..276
Intron 28   :   277..>

```
        accttcctca tcacccttgg gtccctgtct ctctccttcc tgccccttcc          50
        ctctccctgc cccattcctt gcagGGTCCT CAAGCATTCG GTGGACGCCA         100
        CCTTTGAGAA CCAGGGCCCC AGCCCCGGGT ACCGCATGGA GATGTCCATT         150
        TTCTACGTCG TCTACTTTGT GGTGTTCCCC TTCTTCTTTG TCAATATCTT         200
        TGTGGCCTTG ATCATCATCA CCTTCCAGGA GCAAGGGGAC AAGATGATGG         250
        AGGAATACAG CCTGGAGAAA AATGAGgtgc acttccaat tccatctgtc          300
        ctttaaaaac tggggacaca cacaaacttt aaaacacaca caacacccag         350
        gaacccccttt ctaggggtac ctgggggagg aacagaagc attgtcccaa         400
        ccgaatccag tcttcagggc agcccttcat ggagtttcag aggaaacaca         450
        tcatatagtg tatgtatcag tcagtttaga ctaggttat                     489
```

```
//
Submission no   :    27
Intron 28       :    <..253
Exon 29         :    254..418
Intron  29      :    419..>


    tagcccatgc aanaatgggg aaatgncagt gcaagttttg gcagttgntg        50
    acatctcaag caactgtanc tgttgggata agaaagcaat ggtgagaagg       100
    aanagaganc ccaggaatcc tggctggggg caananaggc agagactcaa       150
    gcagaagcac ttgagaaccg cgacgagtta gacagagggt gcccggtgta       200
    cagccacctt cctcctgcct ctgccgctct caccactggc ctctctcccg       250 50
    cagAGGGCCT GCATTGATTT CGCCATCAGT GCCAAGCCGC TGACCCGACA       300
    CATGCCGCAG AACAAGCAGA GCTTCCAGTA CCGCATGTGG CAGTTCGTGG       350
    TGTCTCCGCC TTTCGAGTAC ACGATCATGG CCATGATCGC CCTCAACACC       400
    ATCGTGCTTA TGATGAAGgt aagtgcccca caccagcccc cagcactant       450
    taacccccac ctcgttcctg cctctaccct gataaaatga aaccatttgc       500
    agatttccca ga                                               512
```

```
//
Submission no   :    28
Intron 29       :    156
Exon 30         :    157..267
Intron 30       :    268..>
```

```
        gggtctttcc tgaactgtgc ctcctaccag tgaggttgct cagaccttgc          50
        ctggggctgg agtgttgcct ggagaacagc catgaagctg acctccccac         100
        ttcccacttc ccacccctgc tcgctgaccc ctgctactcc tgcttctttc         150
        ccctagTTCT ATGGGGCTTC TGTGGCTTAT GAAAATGCCC TGCGGGTGTT         200
        CAACATCGCC TTCACCTCCC TCTTCTCTCT GGAATGTGTG CTGAAAGCCA         250
        TGGCTTTTGG GATTCTGgta agtaccacct tggggctaca gctatgggct         300
        tggɟanaaanc ccaaggggga acaatgggtc ctggatgatg gtctcccaac         350
        gtggccccaa gaaccccaac ctcaagggtg gcttcagtat cctgcccagt         400
        ggccacagat c                                                    411
```

```
//
Submission no    :    29
Intron 30        :    <..115
Exon 31          :    116..199
Intron 31        :    200..>


    ctgtcccggg  cactccgctg  atgggcaact  gtgcctctaa  catgcaccgg        50
    ccagcctagg  gggccgggaa  ccaagccctc  tgttggcatc  tctgtcttgt       100
    gggtccccat  tctagAATTA  TTTCCGCGAT  GCCTGGAACA  TCTTCGACTT       150
    TGTGACTGTT  CTGGGCAGCA  TCACCGATAT  CCTCGTGACT  GAGTTTGGGg       200
    taagtctccc  tccagcttct  ctctgggtga  ctctgggctg  gacgaggcag       250
    gcggcagggg  gcgggggagc  ggtcccagag  gcagtgtgtc  ccggaagcca       300
    tagctgcttg  agccagcact  tggccatgac  cagagaggga  gaactggggc       350
    cccggggaca  agggcagccc  ctcaggaggg  cattgtgggg  agatggggt       400
    aacaaagctt  ggctgtaggg                                            420
```

```
//
Submission no  :    30
Intron 31      :    <..148
Exon 32        :    149..265
Intron 32      :    266..>


        ttaatagtgc tttctctctc cctccttatt tggggtctgg cttgcttttt        50
        tcctgttggt tggcttcatg taggggcctc tgtgagtggt gacagctctg       100
        agcctttggg gtgggtggat ggtcacccct cttcctccat ctccccagAA       150
        TAACTTCATC AACCTGAGCT TTCTCCGCCT CTTCCGAGCT GCCCGGCTCA       200
        TCAAACTTCT CCGTCAGGGT TACACCATCC GCATTCTTCT CTGGACCTTT       250
        GTGCAGTCCT TCAAGgtgag tcctcgtccc tgctgctggc ccaggggctg       300
        agaagacagg tgaccctcat gctctggctg aatgtagaag tc              342
```

```
//
Submission no  :   31
Intron 32      :   <..156
Exon 33        :   157..222
Intron 33      :   223..394
Exon 34        :   395..509
Intron 34      :   510..>
```

```
cccccaagaa gaatgcccac caagccctgg aaggactctg gcacgtggca      50
tatggccacc caacccagtg gggcagagca ctgggacaag ggaggaagac     100
tgcagtgcgg ctgagggacc cccagcactc ttcttcattg ccttttttcc     150
caccagGCCC TGCCTTATGT CTGTCTGCTG ATCGCCATGC TCTTCTTCAT     200
CTATGCCATC ATTGGGATGC AGgtgagtgt cgtgtcccta aggttcccag     250
agcctcccaa ggagggcagc caccccttaga aaggggtggg tcagaggagc     300
ctggttcaca gaagcagcca tggaggttga gctgggtttc ccagaagcca     350
ctggaggaat ggcagcccct ggtcgtcacc cwmcaattcc acagGTGTTT     400
GGTAACATTG GCATCGACGT GGAGGACGAG GACAGTGATG AAGATGAGTT     450
CCAAATCACT GAGCACAATA ACTTCCGGAC CTTCTTCCAG GCCCTCATGC     500
TCTCTTCCGg tcagaagggg acctgctctg ataatnctgt ttccgtgggg     550
tggggtgcc                                                   559
```

```
//
Submission no  :    32
Intron 34      :    <..94
Exon 35        :    95..245
Intron 35      :    246..>
Sequence       :    316
```

```
tcagagccat gctcactgtg tgctccactc ctgaggaggc gttggtacca          50
gtcagggctg ggtgtccgag tctctgattt ctccctgtcc tcagGAGTGC          100
CACCGGGGAA GCTTGGCACA ACATCATGCT TTCCTGCCTC AGCGGGAAAC          150
CGTGTGATAA GAACTCTGGC ATCCTGACTC GAGAGTGTGG CAATGAATTT          200
GCTTATTTTT ACTTTGTTTC CTTCATCTTC CTCTGCTCGT TTCTGgtgag          250
tctgtggaca ctgtgagggc cgtctgggct ccctaagcct ggcttccttt          300
cagggagtgg ttctgt                                               316
```

```
//
Submission no   :      33
Intron 35       :      <..211
Exon 36         :      212..339
Intron 36       :      340..>
```

```
gtgtagtgag aactcacctc tccattcccc agtctctttc tgtctctgtc          50
tcatttcctt tccccatctt ctctctatcc ctctctccat ctggggcctc         100
tgtgtctgtc tttgggtctg tctgtccgtc tgactgtctg tatccttctc         150
acttcactca ttcattccct cggtctctgc cccattctct cttggtcccg         200
ggtccccaca gATGCTGAAT CTCTTTGTCG CCGTCATCAT GGACAACTTT         250
GAGTACCTCA CCCGAGACTC CTCCATCCTG GGCCCCCACC ACCTGGATGA         300
GTACGTGCGT GTCTGGGCCG AGTATGACCC CGCAGCTTGg taagaagtca         350
ccccgaatcc tccagccaca atactcacct ctccctggaa ctggaacacg         400
ggctaggcta ggnccccaga ctctggagca ctgaactcct ggggctccta         450
gcagggGtct cacaggttca gtcaggagag aagatataag aatcatcacc         500
cttgcatacc ccagattaaa cacgtagggt gccaacctgc ccaaaccctg         550
gaggactttc tgggaaatga ggaggGcgtc aaccatgaga tgtctgaaga         600
gccctctcct cctacgagtc tctcctgtct ctcactgtga agtctccaga         650
tggtgaggat cgattagcca ggctccagga gaaaccaaca gact              694
```

```
//
Submission no   :    34
Intron 36       :    <..213
Exon 37         :    214..310
Intron 37       :    311..>
```

```
        aagggaggtg cctgcagtcc cgaactcgac tgacatccta cacccctggg        50
        tctccccagt gtctgggaat gtactgggaa ttcacttgtc cccagtctct       100
        cccactcctt gaagccaggg acaccccagc ctcgggcatc atgacctcgt       150
        tgtgtgccca gggagcccgt gtgaacccat tgcctgcact aaccccccttt      200
        cttctccttt cagCGGTCGG ATTCATTATA AGGATATGTA CAGTTTATTA       250
        CGAGTAATAT CTCCCCCTCT CGGCTTAGGC AAGAAATGTC CTCATAGGGT       300
        TGCTTGCAAG gtttgacttc cactaaaacc tgctagcatc catggaatga       350
        gtgtggcttg gggttcttca atatatatat ttcatatata tatatatata       400
        tatctctctc tctctaaaaa aacagagcca tctctctttc ttgcattaaa       450
        ctagaaaact ctcttagcca acag                                   474
```

```
//
Submission no    :    35
Intron 37        :    <..82
Exon 38          :    83..188
Intron 38        :    189..>


        cctgggtagg  ggcgggcgcg  gctcacggga  gacccaggag  ggatgcctgg        50
        gaatgactgc  gcttgccttg  ggttttctgt  agCGGCTTCT  GCGGATGGAC       100
        CTGCCCGTCG  CAGATGACAA  CACCGTCCAC  TTCAATTCCA  CCCTCATGGC       150
        TCTGATCCGC  ACAGCCCTGG  ACATCAAGAT  TGCCAAGGgt  aaggaaggga       200
        caggggcggg  cacagacagg  cgtgacaggg  tggaactggg  gatctcctcc       250
        ctaccccaaa  ctagaggatc  tgctgtcacc  acccggatct  tcattcactc       300
        ttccattcat  tcgttccaca  ggnntttttg  gnnnttggnn  ntttggtgtt       350
        tttttttttt  ttttgagaca  gagtcttgct  ctgttgccca  ggcagcagtg       400
        cggtgacatg  atc                                                  413
```

```
//
Submission no  :   36
Intron 38      :   <..96
Exon 39        :   97..204
Intron 39      :   205..369
Exon 40        :   370..470
Intron 40      :   471..>
```

```
gggtctcgtt  ctcgggagcc  tatggctttg  cagctgaccc  agagtccagc        50
tgacacccag  gcaggcagtc  agggtctgtc  tacaccccca  ttgcagGAGG       100
AGCCGACAAA  CAGCAGATGG  ACGCTGAGCT  GCGGAAGGAG  ATGATGGCGA       150
TTTGGCCCAA  TCTGTCCCAG  AAGACGCTAG  ACCTGCTGGT  CACACCTCAC       200
AAGTgtaaga  gctgagccca  gccctgggat  ccaatccacc  aggacagatg       250
gag ggggagg  gaaagggggag  gcctggggag  agtgttggct  gggctggtat       300
acacagggac  ccaggacaag  gtccccaaag  angcctgccc  ttggtgagct       350
caccgtgtgt  gtcccccagC  CACGGACCTC  ACCGTGGGGA  AGATCTACGC       400
AGCCATGATG  ATCATGGAGT  ACTACCGGCA  GAGCAAGGCC  AAGAAGCTGC       450
AGGCCATGCG  CGAGGAGCAG  gtgcgctgtt  cgccgctctg  gggacatctg       500
ggctgggac  agtggcttgc  atgtcaccac  gggaaccaac  tggaatatga       550
gggtggctga  gcccagggc  aggtccctga  aaagtagggg  ctggtgcaca       600
gcagctcaca  cctgcaatct  cagtgctttg  agaggc                       636
```

```
//
Submission no    :    37
Intron 40        :    <..407
Exon 41          :    408..517
Intron 41        :    518..625
Exon 42          :    626..764
Intron 42        :    765..>
Sequence         :    829
```

```
          gatcttcagg gccatgggag ctgcaggaag gactctggct ttttccccaa          50
          gcaagtggga gccatggagg gttctaagca aaggagggat aggacctgac         100
          tcaagtgctc atgggcgccc tctggtggct cttgtggaac agtggggttg         150
          aaggtaggag cgggagacct gggagaaggt gcctgcagtg agagatgagg         200
          acgcgggacc aggctggggc tatgacttgg gtggaggagt gagaagtggt         250
          ccagttctgc gtggaattgg aagggtctag atggatgaga cctgagagag         300
          tgtgtgtgtg tgtgtgtgtg tatactgggg atgtcgcaat gccttctggg         350
          taccaccgtc caccacccca cccttgtcca cacactgctc tctgccccat         400
          tccccagGAC CGGACACCCC TCATGTTCCA GCGCATGGAG CCCCCGTCCC         450
          CAACGCAGGA AGGGGGACCT GGCCAGAACG CCCTCCCCTC CACCCAGCTG         500
          GACCCAGGAG GAGCCCTgtg agtgtcaccc ctgccaggga ggtggagtgt         550
          gggggtgccg tggtccccac gttctggaag ctgcccaagc gcccactgct         600
          accccggcct ctgtccccca tgcagGATGG CTCACGAAAG CGGCCTCAAG         650
          GAGAGCCCGT CCTGGGTGAC CCAGCGTGCC CAGGAGATGT TCCAGAAGAC         700
          GGGCACATGG AGTCCGGAAC AAGGCCCCCC TACCGACATG CCCAACAGCC         750
          AGCCTAACTC TCAGgtgcct ctgtccccca actccccaat ggctcccagg         800
          gcccgggtgg ttgcggtgga aggaaccat                                829
```

```
//
Submission no   :     38
Intron 42       :     <..219
Exon 43         :     220..333
Intron 43       :     334..>
```

```
      tcactgcaac ctccaccttc cagtctcaag tgattcctcc tgcctcagcc        50
      tcccaagtca ctggattaca ggcgcccacc accatgctca ggtatttttt       100
      tttgtatttt tagtagagac ggggtttcac aatgttggtc aggctggtct       150
      cgaactgctg nccattgtga tctggaggtc aggccccaga gctcatctgg       200
      ctttgccatt cgtccgcagT CCGTGGAGAT GCGAGAGATG GGCAGAGATG       250
      GCTACTCCGA CAGCGAGCAC TACCTCCCCA TGGAAGGCCA GGGCCGGGCT       300
      GCCTCCATGC CCCGCCTCCC TGCAGAGAAC CAGgtgaggg ctttcaccac       350
      tgccctgggg ctggacccct cactctgcac tgggtagggc caggcccccc       400
      cacaagcagc ccagtgcatc ccctcctgcc ggactcaggc ctgggtaggg       450
      actccttcag tctctgaagc agtctgcagg ccccacccac cacctggtca       500
      cacctggagc acctgcagac cctcctccct cacagaggac agagaggaaa       550
      gtgctccccc tggggcagag ggcagtggcc actgcaaaat ggtctctggc       600
      tgccctggtt ggaggctgca gacaggggag gttgtggaar atttgtgggt       650
      gcagcagggt tcaacagggc cagctgagac ctgccacgaa gawcctttga       700
      ggccaggagt ttgagaccag gttgggcaac atagcaaaac cctgtctctt       750
      tttttttttt gagacggagt ttcactcttg ttgccccagg ctggagtgac       800
      a                                                             801
```

```
//
Submission no    :    39
Intron 43        :    <..83
Exon 44          :    84..119
Intron 44        :    120..>
Sequence         :    329


        cctcctcact cttccctctt gcctttatat ttattttctt ctttctgttt         50
        tttctgtgtg caccatccat ggggctgtga cagAGGAGAA GGGGCCGGCC        100
        ACGTGGGAAT AACCTCAGTg tatgtacggc ctgcccaggg cccagcaggc        150
        tccggccccc tcttcctccc cacccchcct ccagggagtc ccgtaatctc        200
        taccggtccc cggaccccac cctttctttg gcaatcgcac cctctcccct        250
        ccatggagcc caatccttgt gtgtggtgtc ctgtgtgtgc cctgacccat        300
        aagcctggtg gggcggccat ccccatcct                              329
```

```
//
Submission no   :   40
Intron 44       :   <..166
Exon 45         :   167..353
Intron 45       :   354..>
```

```
gatcaggggg agccaaggcc ccatggcatc ccctggcccc tgccccagga      50
tggtcacacc gcagtcaccg aaggccacca ccaggctgcc acaatggggc     100
aggaaggacc gggaccactt ggtgctagct gctgacccca gcccaccggc     150
ctgtcccctc cccagACCA TCTCAGACAC CAGCCCCATG AAGCGTTCAG       200
CCTCCGTGCT GGGCCCCAAG GCCCGACGCC TGGACGATTA CTCGCTGGAG      250
CGGGTCCCGC CCGAGGAGAA CCAGCGGCAC CACCAGCGGC GCCGCGACCG      300
CAGCCACCGC GCCTCTGAGC GCTCCCTGGG CCGCTACACC GATGTGGACA      350
CAGgtgggca gccctgtggt gctcagggac aagcagaaca gaggagagga      400
gaggggagga gaaggcaggg cggaggagac actaaggaag aagaaaggga      450
gaggcctcca tggagagggg acagagcggg ccaggcagcg gctgcaggaa      500
cctgggtact accccctccc cccaacccac tgacctgcct cggttcaggg      550
gatc                                                        554
```

```
//
Submission no    :    41
Intron 45        :    <..31
Exon 46          :    32..285
Intron 46        :    286..>
```

```
ctgtgtgctg tctgaccctc acccggccca gGCTTGGGGA CAGACCTGAG        50
CATGACCACC CAATCCGGGG ACCTGCCGTC GAAGGAGCGG GACCAGGAGC       100
GGGGCCGGCC CAAGGATCGG AAGCATCGAC AGCACCACCA CCACCACCAC       150
CACCACCACC ATCCCCCGCC CCCCGACAAG GACCGCTATG CCCAGGAACG       200
GCCGGACCAC GGCCGGGCAC GGGCTCGGGA CCAGCGCTGG TCCCGCTCGC       250
CCAGCGAGGG CCGAGAGCAC ATGGCGCACC GGCAGgtggg tgcggctgca       300
agtgacccca ggctgggctc ggccgggagg cggggaggag agaaggggat       350
accccatcca acagccactc taggcaaagg tccccggatc ccggctgtga       400
ccacctccca tcctgccccc aagccaccgg ggtgcccggc ggccggagcg       450
gagcacggat c                                                461
```

```
//
Submission no    :    42
Intron 46        :    <..279
Exon 47          :    280..>
Stop codon       :    280..282
UTR 3'           :    283..>


         tttctcattt ctcttttcac ttttgttgtg ttggtttccg actcctcccc            50
         tccctgtctc actccccctc ctccctcccc tcctccctgt ggctgttgct           100
         tttttccatt caatgtcctg tgtcccccct ctcctcctcc tcctcctcct           150
         ccccctcctc cctctcctcc cggcccctct cccttcgctc ccctcatctt           200
         cctcccaatc ccgtgtctcc tttgattttg ttgtatcttt tttttgatt            250
         tcctttgttt caattttcgt gtagggcagT AGTTCCGTAA GTGGAAGCCC           300
         AGCCCCCTCA ACATCTGGTA CCAGCACTCC GCGGCGGGGC CGCCGCCAGC           350
         TCCCCCAGAC CCCCTCCACC CCCCGGCCAC ACGTGTCCTA TTCCCCTGTG           400
         ATCCGTAAGG CCGGCGGCTC GGGGCCCCCG CAGCAGCAGC AGCAGCAGCA           450
         GCAGCAGCAG CAGGCGGTGG CCAGGCCGGC CGGGCGGCCA CCAGCGGCCC           500
         TCGGAGGTAC CCAGGCCCCA CGGCCGAGCC TCTGGCCGGA GATCGGCGCC           550
         CACGGGGGGC CACAGCAGCG GCCGCACGCC CAGGATGGAG AGGCGGGTCC           600
         AGGCCCGGCC CGGAGCGAGT CTCCAGGGCC TGGTCGACAC GGCGGGGCCC           650
         GGCTGGCGGC AGTC                                                  664
```
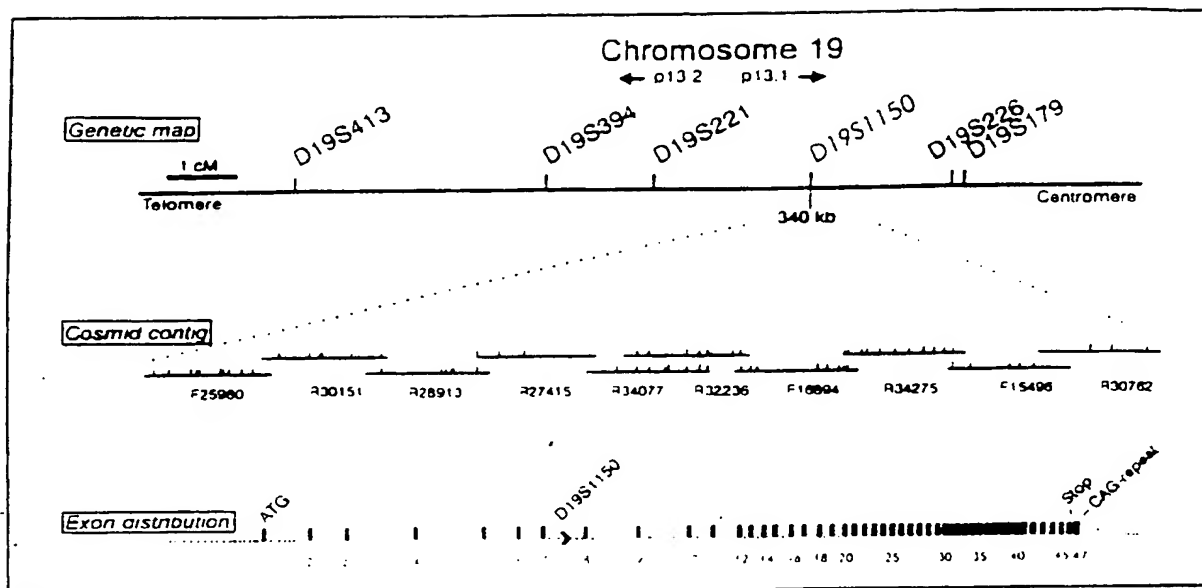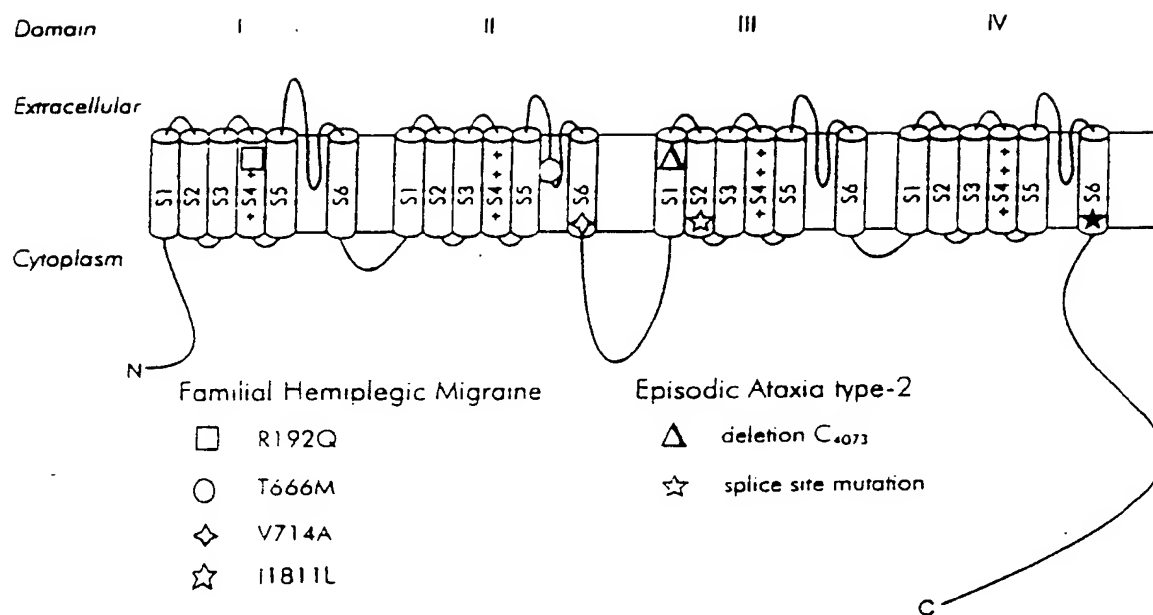
Fig. 2

Fig. 3

Figure 4

```
atggcccgcttcggagacgagatgccggcccgctacggggggaggaggctccggggcagcc        60
 M   A   R   F   G   D   E   M   P   A   R   Y   G   G   G   G   S   G   A   A       20
gccggggtggtcgtgggcagcggaggcgggcgaggagccggggggcagccggcagggcggg       120
 A   G   V   V   V   G   S   G   G   G   R   G   A   G   G   S   R   Q   G   G       40
cagcccgggggcgcaaaggatgtacaagcagtcaatggcgcagagagcgcggaccatggca       180
 Q   P   G   A   Q   R   M   Y   K   Q   S   M   A   Q   R   A   R   T   M   A       60
ctctacaaccccatccccgtccgacagaactgcctcacggttaaccggtctctcttcctc       240
 L   Y   N   P   I   P   V   R   Q   N   C   L   T   V   N   R   S   L   F   L       80
ttcagcgaagacaacgtggtgagaaaatacgccaaaaagatcaccgaatggcctccctt       300
 F   S   E   D   N   V   V   R   K   Y   A   K   K   I   T   E   W   P   P   F       100
gaatatatgattttagccaccatcatagcgaattgcatcgtcctcgcactggagcagcat       360
 E   Y   M   I   L   A   T   I   I   A   N   C   I   V   L   A   L   E   Q   H       120
ctgcctgatgatgacaagaccccgatgtctgaacggctggatgacacagaaccatacttc       420
 L   P   D   D   D   K   T   P   M   S   E   R   L   D   D   T   E   P   Y   F       140
attggaattttttgtttcgaggctggaattaaaatcattgcccttgggtttgccttccac       480
 I   G   I   F   C   F   E   A   G   I   K   I   I   A   L   G   F   A   F   H       160
aaaggctcctacttgaggaatggctggaatgtcatggactttgtggtggtgctaacgggc       540
 K   G   S   Y   L   R   N   G   W   N   V   M   D   F   V   V   V   L   T   G       180
atcttggcgacagttgggacggagtttgacctacggacgctgagggcagttcgagtgctg       600
 I   L   A   T   V   G   T   E   F   D   L   R   T   L   R   A   V   R   V   L       200
cggccgctcaagctggtgtctggaatcccaagtttacaagtcgtcctgaagtcgatcatg       660
 R   P   L   K   L   V   S   G   I   P   S   L   Q   V   V   L   K   S   I   M       220
aaggcgatgatcccctttgctgcagatcggcctcctcctattttttgcaatccttatTtt       720
 K   A   M   I   P   L   L   Q   I   G   L   L   L   F   A   I   L   I   F       240
gcaatcataggggttagaattttatatgggaaaatttcataccacctgctttgaagagggg       780
 A   I   I   G   L   E   F   Y   M   G   K   F   H   T   T   C   F   E   E   G       260
acagatgacattcagggtgagtctccggctccatgtgggacagaagagcccgcccgcacc       840
 T   D   D   I   Q   G   E   S   P   A   P   C   G   T   E   E   P   A   R   T       280
tgccccaatgggaccaaatgtcagccctactgggaagggcccaacaacgggatcactcag       900
 C   P   N   G   T   K   C   Q   P   Y   W   E   G   P   N   N   G   I   T   Q       300
ttcgacaacatcctgtttgcagtgctgactgtttttccagtgcataaccatggaagggtgg       960
 F   D   N   I   L   F   A   V   L   T   V   F   Q   C   I   T   M   E   G   W       320
actgatctcctctacaatagcaacgatgcctcagggaacacttggaactggttgtacttc      1020
 T   D   L   L   Y   N   S   N   D   A   S   G   N   T   W   N   W   L   Y   F       340
atcccctcatcatcatcggctcctttttttatgctgaaccttgtgctgggtgtgctgtca      1080
 I   P   L   I   I   I   G   S   F   F   M   L   N   L   V   L   G   V   L   S       360
ggggagtttgccaaagaaagggaacgggtggagaaccggcgggcttttctgaagctgagg      1140
 G   E   F   A   K   E   R   V   E   N   R   R   A   F   L   K   L   R       380
cggcaacaacagattgaacgtgagctcaatgggtacatggaatggatctcaaaagcagaa      1200
 R   Q   Q   Q   I   E   R   E   L   N   G   Y   M   E   W   I   S   K   A   E       400
gaggtgatcctcgccgaggatgaaactgacggggagcagaggcatccctttgatggagct      1260
 E   V   I   L   A   E   D   E   T   D   G   E   Q   R   H   P   F   D   G   A       420
ctgcggagaaccaccataaagaaaagcaagacagatttgctcaaccccgaagaggctgag      1320
 L   R   R   T   T   I   K   K   S   K   T   D   L   L   N   P   E   E   A   E       440
gatcagctggctgatatagcctctgtggggttctcccttcgcccgagccagcattaaaagt      1380
 D   Q   L   A   D   I   A   S   V   G   S   P   F   A   R   A   S   I   K   S       460
gccaagctggagaactcgacctttttttcacaaaaaggagaggaggatgcgttctacatc      1440
 A   K   L   E   N   S   T   F   F   H   K   K   E   R   R   M   R   F   Y   I       480
cgccgcatggtcaaaactcaggccttctactggactgtactcagtttggtagctctcaac      1500
 R   R   M   V   K   T   Q   A   F   Y   W   T   V   L   S   L   V   A   L   N       500
acgctgtgtgttgctattgttcactacaaccagcccgagtggctctccgacttcctttac      1560
 T   L   C   V   A   I   V   H   Y   N   Q   P   E   W   L   S   D   F   L   Y       520
tatgcagaattcatttttcttaggactcttatgtccgaaatgtttataaaaatgtacggg      1620
 Y   A   E   F   I   F   L   G   L   F   M   S   E   M   F   I   K   M   Y   G       540
cttgggacgcggccttacttccactcttccttcaactgctttgactgtggggttatcatt      1680
 L   G   T   R   P   Y   F   H   S   S   F   N   C   F   D   C   G   V   I   I       560
gggagcatcttcgaggtcatctgggctgtcataaaacctggcacatcctttggaatcagc      1740
```

```
 G   S   I   F   E   V   I   W   A   V   I   K   P   G   T   S   F   G   I   S        580
gtgttacgagccctcaggttattgcgtattttcaaagtcacaaagtactgggcatctctc                    1800
 V   L   R   A   L   R   L   L   R   I   F   K   V   T   K   Y   W   A   S   L        600
agaaacctggtcgtctctctcctcaactccatgaagtccatcatcagcctgttgtttctc                    1860
 R   N   L   V   V   S   L   L   N   S   M   K   S   I   I   S   L   L   F   L        620
cttttcctgttcattgtcgtcttcgccctttttgggaatgcaactcttcggcggccagttt                    1920
 L   F   L   F   I   V   V   F   A   L   L   G   M   Q   L   F   G   G   Q   F        640
aatttcgatgaagggactcctcccaccaacttcgatactttttccagcagcaataatgacg                    1980
 N   F   D   E   G   T   P   P   T   N   F   D   T   F   P   A   A   I   M   T        660
gtgtttcagatcctgacgggcgaagactggaacgaggtcatgtacgacgggatcaagtct                    2040
 V   F   Q   I   L   T   G   E   D   W   N   E   V   M   Y   D   G   I   K   S        680
cagggggggcgtgcagggcggcatggtgttctccatctatttcattgtactgacgctcttt                    2100
 Q   G   G   V   Q   G   G   M   V   F   S   I   Y   F   I   V   L   T   L   F        700
gggaactacaccctcctgaatgtgttcttggccatcgctgtggacaatctggccaacgcc                    2160
 G   N   Y   T   L   L   N   V   F   L   A   I   A   V   D   N   L   A   N   A        720
caggagctcaccaaggacgagcaagaggaagaagaagcagcgaaccagaaacttgcccta                    2220
 Q   E   L   T   K   D   E   Q   E   E   E   E   A   A   N   Q   K   L   A   L        740
cagaaagccaaggaggtggcagaagtgagtcctctgtccgcggccaacatgtctatagct                    2280
 Q   K   A   K   E   V   A   E   V   S   P   L   S   A   A   N   M   S   I   A        760
gtgaaagagcaacagaagaatcaaaagccagccaagtccgtgtggggagcagcggaccagt                    2340
 V   K   E   Q   Q   K   N   Q   K   P   A   K   S   V   W   E   Q   R   T   S        780
gagatgcgaaagcagaacttgctggccagccgggaggccctgtataacgaaatggacccg                    2400
 E   M   R   K   Q   N   L   L   A   S   R   E   A   L   Y   N   E   M   D   P        800
gacgagcgctggaaggctgcctacacgcggcacctgcggccagacatgaagacgcacttg                    2460
 D   E   R   W   K   A   A   Y   T   R   H   L   R   P   D   M   K   T   H   L        820
gaccggccgctggtggtggacccgcaggagaaccgcaacaacaacaccaacaagagccgg                    2520
 D   R   P   L   V   V   D   P   Q   E   N   R   N   N   N   T   N   K   S   R        840
gcggccgagcccaccgtggaccagcgcctcggccagcagcgcgccgaggacttcctcagg                    2580
 A   A   E   P   T   V   D   Q   R   L   G   Q   Q   R   A   E   D   F   L   R        860
aaacaggcccgctaccacgatcgggcccgggaccccagcggctcggcgggcctggacgca                    2640
 K   Q   A   R   Y   H   D   R   A   R   D   P   S   G   S   A   G   L   D   A        880
cggaggccctgggcgggaagccaggaggccgagctgagccgggaggacccctacggccgc                    2700
 R   R   P   W   A   G   S   Q   E   A   E   L   S   R   E   D   P   Y   G   R        900
gagtcggaccaccacgcccgggagggcagcctggagcaacccgggttctgggagggcgag                    2760
 E   S   D   H   H   A   R   E   G   S   L   E   Q   P   G   F   W   E   G   E        920
gccgagcgaggcaaggccgggggaccccaccggaggcacgtgcaccggcaggggggcagc                    2820
 A   E   R   G   K   A   G   D   P   H   R   R   H   V   H   R   Q   G   G   S        940
agggagagccgcagcgggtccccgcgcacgggcgcggacggggagcatcgacgtcatcgc                    2880
 R   E   S   R   S   G   S   P   R   T   G   A   D   G   E   H   R   R   H   R        960
gcgcaccgcaggcccgggggaggagggtccggaggacaaggcggagcggagggcgcggcac                    2940
 A   H   R   R   P   G   E   E   G   P   E   D   K   A   E   R   R   A   R   H        980
cgcgagggcagccggccggcccggggcggcgagggcgagggcgagggtcccgacggggggc                    3000
 R   E   G   S   R   P   A   R   G   G   E   G   E   G   E   G   P   D   G   G       1000
gagcgcaggagaaggcaccggcatggcgctccagccacgtacgaggggacgcgcggagg                    3060
 E   R   R   R   R   H   R   H   G   A   P   A   T   Y   E   G   D   A   R   R       1020
gaggacaaggagcggaggcatcgggaggaggaaagagaaccagggctccggggtccctgtg                    3120
 E   D   K   E   R   R   H   R   R   R   K   E   N   Q   G   S   G   V   P   V       1040
tcgggccccaacctgtcaaccacccggccaatccagcaggacctgggccgccaagaccca                    3180
 S   G   P   N   L   S   T   T   R   P   I   Q   Q   D   L   G   R   Q   D   P       1060
cccctggcagaggatattgacaacatgaagaacaacaagctggccaccgcggagtcggcc                    3240
 P   L   A   E   D   I   D   N   M   K   N   N   K   L   A   T   A   E   S   A       1080
gctccccacggcagccttggccacgccggcctgccccagagcccagccaagatgggaaac                    3300
 A   P   H   G   S   L   G   H   A   G   L   P   Q   S   P   A   K   M   G   N       1100
agcaccgaccccggccccatgctggccatccctgccatggccaccaacccccagaacgcc                    3360
 S   T   D   P   G   P   M   L   A   I   P   A   M   A   T   N   P   Q   N   A       1120
gccagccgccggacgcccaacaacccggggaacccatccaatcccggcccccccaagacc                    3420
 A   S   R   R   T   P   N   N   P   G   N   P   S   N   P   G   P   P   K   T       1140
```

```
cccgagaatagccttatcgtcaccaaccccagcggcacccagaccaattcagctaagact     3480
 P   E   N   S   L   I   V   T   N   P   S   G   T   Q   T   N   S   A   K   T   1160
gccaggaaacccgaccacaccacagtggacatccccccagcctgcccaccccccctcaac     3540
 A   R   K   P   D   H   T   T   V   D   I   P   P   A   C   P   P   P   L   N   1180
cacaccgtcgtacaagtgaacaaaaacgccaacccagacccactgccaaaaaaagaggaa     3600
 H   T   V   V   Q   V   N   K   N   A   N   P   D   P   L   P   K   K   E   E   1200
gagaagaaggaggaggaggaagaagacgaccgtggggaagacggccctaagccaatgcct     3660
 E   K   K   E   E   E   E   D   D   R   G   E   D   G   P   K   P   M   P   1220
ccctatagctccatgttcatcctgtccacgaccaacccccttcgccgcctgtgccattac     3720
 P   Y   S   S   M   F   I   L   S   T   T   N   P   L   R   R   L   C   H   Y   1240
atcctgaacctgcgctactttgagatgtgcatcctcatggtcattgccatgagcagcatc     3780
 I   L   N   L   R   Y   F   E   M   C   I   L   M   V   I   A   M   S   S   I   1260
gccctggccgccgaggaccctgtgcagcccaacgcacctcggaacaacgtgctgcgatac     3840
 A   L   A   A   E   D   P   V   Q   P   N   A   P   R   N   N   V   L   R   Y   1280
tttgactacgtttttacaggcgtctttacctttgagatggtgatcaagatgattgacctg     3900
 F   D   Y   V   F   T   G   V   F   T   F   E   M   V   I   K   M   I   D   L   1300
gggctcgtcctgcatcagggtgcctacttccgtgacctctggaatattctcgacttcata     3960
 G   L   V   L   H   Q   G   A   Y   F   R   D   L   W   N   I   L   D   F   I   1320
gtggtcagtggggccctggtagcctttgccttcactggcaatagcaaaggaaaagacatc     4020
 V   V   S   G   A   L   V   A   F   A   F   T   G   N   S   K   G   K   D   I   1340
aacacgattaaatccctccgagtcctccgggtgctacgacctcttaaaaccatcaagcgg     4080
 N   T   I   K   S   L   R   V   L   R   V   L   R   P   L   K   T   I   K   R   1360
ctgccaaagctcaaggctgtgtttgactgtgtggtgaactcacttaaaaacgtcttcaac     4140
 L   P   K   L   K   A   V   F   D   C   V   V   N   S   L   K   N   V   F   N   1380
atcctcatcgtctacatgctattcatgttcatcttcgccgtggtggctgtgcagctcttc     4200
 I   L   I   V   Y   M   L   F   M   F   I   F   A   V   V   A   V   Q   L   F   1400
aagggaaaattcttccactgcactgacgagtccaaagagtttgagaaagattgtcgaggc     4260
 K   G   K   F   F   H   C   T   D   E   S   K   E   F   E   K   D   C   R   G   1420
aaatacctcctctacgagaagaatgaggtgaaggcgcgagaccggggagtggaagaagtat     4320
 K   Y   L   L   Y   E   K   N   E   V   K   A   R   D   R   E   W   K   K   Y   1440
gaattccattacgacaatgtgctgtgggctctgctgaccctcttcaccgtgtccacggca     4380
 E   F   H   Y   D   N   V   L   W   A   L   L   T   L   F   T   V   S   T   A   1460
gaaggctggccacaggtcctcaagcattcggtggacgccacctttgagaaccagggcccc     4440
 E   G   W   P   Q   V   L   K   H   S   V   D   A   T   F   E   N   Q   G   P   1480
agccccgggtaccgcatggagatgtccattttctacgtcgtctactttgtggtgttcccc     4500
 S   P   G   Y   R   M   E   M   S   I   F   Y   V   V   Y   F   V   V   F   P   1500
ttcttctttgtcaatatctttgtggccttgatcatcatcaccttccaggagcaaggggac     4560
 F   F   F   V   N   I   F   V   A   L   I   I   I   T   F   Q   E   Q   G   D   1520
aagatgatggaggaatacagcctggagaaaaatgagagggcctgcattgatttcgccatc     4620
 K   M   M   E   E   Y   S   L   E   K   N   E   R   A   C   I   D   F   A   I   1540
agtgccaagccgctgacccgacacatgccgcagaacaagcagagcttccagtaccgcatg     4680
 S   A   K   P   L   T   R   H   M   P   Q   N   K   Q   S   F   Q   Y   R   M   1560
tggcagttcgtggtgtctccgcctttcgagtacacgatcatggccatgatcgccctcaac     4740
 W   Q   F   V   V   S   P   P   F   E   Y   T   I   M   A   M   I   A   L   N   1580
accatcgtgcttatgatgaagttctatggggcttctgtggcttatgaaaatgccctgcgg     4800
 T   I   V   L   M   M   K   F   Y   G   A   S   V   A   Y   E   N   A   L   R   1600
gtgttcaacatcgccttcacctccctcttctctctggaatgtgtgctgaaagccatggct     4860
 V   F   N   I   A   F   T   S   L   F   S   L   E   C   V   L   K   A   M   A   1620
tttgggattctgaattatttccgcgatgcctggaacatcttcgactttgtgactgttctg     4920
 F   G   I   L   N   Y   F   R   D   A   W   N   I   F   D   F   V   T   V   L   1640
ggcagcatcaccgatatcctcgtgactgagtttgggaataacttcatcaacctgagcttt     4980
 G   S   I   T   D   I   L   V   T   E   F   G   N   N   F   I   N   L   S   F   1660
ctccgcctcttccgagctgcccggctcatcaaacttctccgtcagggttacaccatccgc     5040
 L   R   L   F   R   A   A   R   L   I   K   L   L   R   Q   G   Y   T   I   R   1680
attcttctctggacctttgtgcagtccttcaaggccctgccttatgtctgtctgctgatc     5100
 I   L   L   W   T   F   V   Q   S   F   K   A   L   P   Y   V   C   L   L   I   1700
gccatgctcttcttcatctatgccatcattgggatgcaggtgtttggtaacattggcatc     5160
```

```
 A   M   L   F   F   I   Y   A   I   I   G   M   Q   V   F   G   N   I   G   I         1720
gacgtggaggacgaggacagtgatgaagatgagttccaaatcactgagcacaataacttc                       5220
 D   V   E   D   E   D   S   D   E   D   E   F   Q   I   T   E   H   N   N   F         1740
cggaccttcttccaggccctcatgcttctcttccggagtgccaccggggaagcttggcac                       5280
 R   T   F   F   Q   A   L   M   L   L   F   R   S   A   T   G   E   A   W   H         1760
aacatcatgctttcctgcctcagcgggaaaccgtgtgataagaactctggcatcctgact                       5340
 N   I   M   L   S   C   L   S   G   K   P   C   D   K   N   S   G   I   L   T         1780
cgagagtgtggcaatgaatttgcttattttactttgtttccttcatcttcctctgctcg                        5400
 R   E   C   G   N   E   F   A   Y   F   Y   F   V   S   F   I   F   L   C   S         1800
tttctgatgctgaatctctttgtcgccgtcatcatggacaactttgagtacctcacccga                       5460
 F   L   M   L   N   L   F   V   A   V   I   M   D   N   F   E   Y   L   T   R         1820
gactcctccatcctgggcccccaccacctggatgagtacgtgcgtgtctgggccgagtat                       5520
 D   S   S   I   L   G   P   H   H   L   D   E   Y   V   R   V   W   A   E   Y         1840
gaccccgcagcttgcggtcggattcattataaggatatgtacagtttattacgagtaata                       5580
 D   P   A   A   C   G   R   I   H   Y   K   D   M   Y   S   L   L   R   V   I         1860
tctccccctctcggcttaggcaagaaatgtcctcatagggttgcttgcaagcggcttctg                       5640
 S   P   P   L   G   L   G   K   K   C   P   H   R   V   A   C   K   R   L   L         1880
cggatggacctgcccgtcgcagatgacaacaccgtccacttcaattccaccctcatggct                       5700
 R   M   D   L   P   V   A   D   D   N   T   V   H   F   N   S   T   L   M   A         1900
ctgatccgcacagccctggacatcaagattgccaagggaggagccgacaaacagcagatg                       5760
 L   I   R   T   A   L   D   I   K   I   A   K   G   G   A   D   K   Q   Q   M         1920
gacgctgagctgcggaaggagatgatggcgatttggcccaatctgtcccagaagacgcta                       5820
 D   A   E   L   R   K   E   M   M   A   I   W   P   N   L   S   Q   K   T   L         1940
gacctgctggtcacacctcacaagtccacggacctcaccgtggggaagatctacgcagcc                       5880
 D   L   L   V   T   P   H   K   S   T   D   L   T   V   G   K   I   Y   A   A         1960
atgatgatcatggagtactaccggcagagcaaggccaagaagctgcaggccatgcgcgag                       5940
 M   M   I   M   E   Y   Y   R   Q   S   K   A   K   K   L   Q   A   M   R   E         1980
gagcaggaccggacacccctcatgttccagcgcatggagcccccgtccccaacgcaggaa                       6000
 E   Q   D   R   T   P   L   M   F   Q   R   M   E   P   P   S   P   T   Q   E         2000
gggggacctggccagaacgccctcccctccacccagctggacccaggaggagccctgatg                       6060
 G   G   P   G   Q   N   A   L   P   S   T   Q   L   D   P   G   G   A   L   M         2020
gctcacgaaagcggcctcaaggagagcccgtcctgggtgacccagcgtgcccaggagatg                       6120
 A   H   E   S   G   L   K   E   S   P   S   W   V   T   Q   R   A   Q   E   M         2040
ttccagaagacgggcacatggagtccggaacaaggccccccctaccgacatgcccaacagc                      6180
 F   Q   K   T   G   T   W   S   P   E   Q   G   P   P   T   D   M   P   N   S         2060
cagcctaactctcagtccgtggagatgcgagagatgggcagagatggctactccgacagc                       6240
 Q   P   N   S   Q   S   V   E   M   R   E   M   G   R   D   G   Y   S   D   S         2080
gagcactacctccccatggaaggccagggccgggctgcctccatgccccgcctccctgca                       6300
 E   H   Y   L   P   M   E   G   Q   G   R   A   A   S   M   P   R   L   P   A         2100
gagaaccagaggagaaggggccggccacgtgggaataacctcagtaccatctcagacacc                       6360
 E   N   Q   R   R   R   G   R   P   R   G   N   N   L   S   T   I   S   D   T         2120
agccccatgaagcgttcagcctccgtgctgggccccaaggcccgacgcctggacgattac                       6420
 S   P   M   K   R   S   A   S   V   L   G   P   K   A   R   R   L   D   D   Y         2140
tcgctggagcgggtcccgcccgaggagaaccagcggcgccaccagcggcgccgcgaccgc                       6480
 S   L   E   R   V   P   P   E   E   N   Q   R   H   H   Q   R   R   R   D   R         2160
agccaccgcgcctctgagcgctccctgggccgctacaccgatgtggacacaggcttgggg                       6540
 S   H   R   A   S   E   R   S   L   G   R   Y   T   D   V   D   T   G   L   G         2180
acagacctgagcatgaccacccaatccggggacctgccgtcgaaggagcgggaccaggag                       6600
 T   D   L   S   M   T   T   Q   S   G   D   L   P   S   K   E   R   D   Q   E         2200
cggggccggcccaaggatcggaagcatcgacagcaccaccaccaccaccaccaccaccac                       6660
 R   G   R   P   K   D   R   K   H   R   Q   H   H   H   H   H   H   H   H   H         2220
catcccccgccccccgacaaggaccgctatgcccaggaacggccggaccacggccgggca                       6720
 H   P   P   P   P   D   K   D   R   Y   A   Q   E   R   P   D   H   G   R   A         2240
cgggctcgggaccagcgctggtcccgctcgcccagcgagggccgagagcacatggcgcac                       6780
 R   A   R   D   Q   R   W   S   R   S   P   S   E   G   R   E   H   M   A   H         2260
cggcagtag                                                                          6786
 R   Q   *                                                                          2282
```